



Understanding 'what works': Evidence based regional policy making in England.

BAXENDALE, Jessica Kate.

Available from the Sheffield Hallam University Research Archive (SHURA) at:

<http://shura.shu.ac.uk/19331/>

A Sheffield Hallam University thesis

This thesis is protected by copyright which belongs to the author.

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.

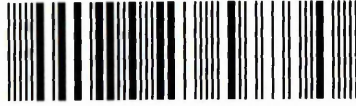
When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.

Please visit <http://shura.shu.ac.uk/19331/> and <http://shura.shu.ac.uk/information.html> for further details about copyright and re-use permissions.

Sheffield Hallam University
Learning and Information Services
Adsetts Centre, City Campus
Sheffield S1 1WD

28102

102 144 998 9



Sheffield Hallam University
Learning and Information Services
Adsetts Centre, City Campus
Sheffield S1 1WD

REFERENCE

ProQuest Number: 10694212

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 10694212

Published by ProQuest LLC (2017). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

Understanding 'What Works'?
Evidence Based Regional Policy Making in England

Jessica Kate Baxendale

A thesis submitted in partial fulfilment of the requirements of
Sheffield Hallam University
for the degree of Doctor of Philosophy

September 2015

Abstract

AIM: This thesis tests the applicability of, and the factors which determine the success of, Evidence Based *Regional* Policy Making in England.

CONTEXT: The concept of Evidence Based Policy Making (EBPM) has been mobilised by the New Labour and Coalition governments through the 'what works' agenda. A significant contention underlying current debates about EBPM, and in turn debates about the role of evaluation evidence, has focused on the credibility of study findings. This has led to a call for the extension of approaches more closely aligned with Evidence Based Medicine (EBM) to wider social policy, such as the use of experimentation, econometrics and economic evaluation.

OBJECTIVES: Firstly, to take an interdisciplinary approach across the health and regional policy sectors to investigate what types of evidence are used and the role of research credibility. Secondly, to investigate the generation, communication and use of evaluation evidence within the Regional Development Agencies to understand, not only the way in which evidence was incorporated into regional policy making processes, but the role of other factors besides evidence. Finally, to critically analyse the use of a knowledge translation tool as a strategy to increase the uptake of evidence.

METHODS: A comparative analysis of the methodological guidelines for evaluation evidence across the health and regional policy sectors was undertaken. A mixed-methods approach was then taken to explore the views of expert stakeholder groups involved in RDA policy evaluation. This included an online survey, the development of a knowledge translation tool and an online workshop to test the applicability of such a tool to the regional policy context. Ninety-five policy makers and analysts contributed towards the research.

CONCLUSIONS: A central finding of the thesis is the need for a more nuanced approach to the generation and use of evidence. This is in contrast to imposing a quality criteria specific to one type of study design (e.g. experimental methods) or allowing for cherry-picked and unsystematic evidence use within policy making processes.

It is also argued that the development of a knowledge translation tool, operationalised through an evaluation and monitoring framework from the start of an organisation's existence, may facilitate the collection of more appropriate, decision-relevant data linked to an underlying programme theory. This would enable a tangible understanding of how data is to be aggregated, highlight any knowledge gaps and facilitate data-linking to other intelligence sources. The inclusion of policy makers early in the research process may also enable the generation of problem-driven evidence and shape an understanding of how such evidence supports decision making.

Acknowledgements

Foremost, I would like to thank Professor Peter Wells and Ian Wilson. I consider myself very fortunate to have had such knowledgeable and unfailingly supportive PhD supervisors. Under your guidance I have matured and developed as a researcher and for that I will always be grateful. Particular thanks go to Peter for inspiring me to explore the political nature of evidence and decision making and to Ian for your methodological insight. I would also like to express my appreciation to the ESRC for providing me with a studentship so that I could complete this PhD.

Heartfelt thanks go to Dr Sue Richardson for your encouragement of me as a researcher over the years. I am indebted to Sue and to Mike Chadwick for providing excellent feedback on an earlier version of this thesis, to Brian Bennett for your feedback on Chapter 4, to Ruth Clark for your technical support with the online workshop, and to Professor Paul Lawless for your insight at the start of my PhD journey.

Thanks to all of the research participants who were generous with their time in a way that I can never repay. Particular thanks to Dr Carol Murray, William Rossiter and Carol Candler for assisting me in recruitment. I would also like to thank all of the staff at the Centre for Regional Economic and Social Research and particularly to Ryan Powell for the PhD forums.

During the course of this PhD process my friends and colleagues have cheered me on and sustained patient interest while I discussed the challenges of my work. Thank you Ruth, Sarah, Holly, Helen, Sarah, Becky, Vicky, Becky, Tracy, Kate, Aimee, Jen and Ellen. Thanks also to Hayley. Special thanks go to my family for supporting me including Lesley, Chris, Ben, Ellie, Jo and Matthew. To my grandparents Barbara, Jack, Audrey, Cyril and to the honorary members John and Jackie. I am also beyond grateful to my parents, Helen and Mick. You have instilled in me a belief that I can do anything I put my mind to and have supported me unwaveringly throughout this process despite facing your own challenging times.

Thank you to my little boy, Oscar. Your arrival has brought me so much joy and time with you has been a welcome distraction. Finally, but most important of all, thanks are due to my husband, Alex, for quietly supporting me behind the scenes and giving me the space to write. I wouldn't have been able to do it if it wasn't for you.

Table of Contents

Abstract	ii
Acknowledgements.....	iii
Candidate's Statement.....	iv
Table of Contents	v
Figures	xii
Tables	xiii
Chapter 1.....	1
Introduction	1
1.1 Introduction	1
1.2 Context: Regional policy in England.....	2
1.2.1 Spatially unbalanced growth in England.....	3
1.3 A brief history of regional policy in England: 1928 to the present day	5
1.3.1 Aims.....	5
1.3.2 Institutions and participants	11
1.4 Regional policy evaluation in England: the current situation in 2015	15
1.5 Defining the research problem	18
1.5.1 The conceptual framework: Evidence Based Policy Making	18
1.5.2 Gap in knowledge: Evidence Based Regional Policy Making	19
1.6 Research aims	21
1.6.1 Adapting to the research context	22
1.6.2 My place vis-à-vis the research.....	23
1.6.3 Thesis overview.....	24
Chapter 2.....	26
Review of the EBPM literature.....	26

2.1 Introduction	26
2.2 Literature review methods.....	26
2.3 Debate 1: Evidence types and the role of research credibility	27
2.4 Debate 2: The way in which evidence is incorporated into the policy making process	32
2.5 Debate 3: Other factors besides evidence which affect the way policy is made .	39
2.6 Debate 4: The role of knowledge translation tools and decision support	41
2.7 Conclusion	45
Chapter 3	46
Methodology	46
3.1 Introduction	46
3.2 Rationale for the methodology	46
3.2.1 Practical considerations	47
3.2.2 Philosophical considerations.....	47
3.3 Methodological framework	48
3.3.1 Sampling framework	50
3.4 Initial scoping interviews.....	53
3.5 Phase 1: Comparative literature review	53
3.6 Phase 2: Online survey	55
3.6.1 Data collection	56
3.6.2 Data analysis	62
3.7 Phase 3: Decision support tool	64
3.7.1 Developing the tool.....	65
3.8 Phase 4: Online workshop.....	75
3.8.1 Background	76
3.8.2 Data collection	79
3.8.3 Data analysis	98

3.9 Conclusion	98
Chapter 4.....	100
Health Policy & Evidence Based Medicine	100
4.1 Introduction	100
4.1.1 Theoretical background	100
4.1.2 Scope: Health policy in England	100
4.2 Evidence Based Medicine (EBM).....	101
4.2.1 A biomedical model of health	101
4.3 Evidence types and the role of research credibility.....	102
4.4 The way in which evidence is incorporated into the policy making process.....	105
4.4.1 Health policy in England and resource allocation.....	105
4.4.2 The NICE evaluation model.....	106
4.4.3 Study designs and methods for NICE economic evaluation	108
4.5 Other factors besides evidence which affect the way policy is made	112
4.6 The extension of EBM to health and social care.....	114
4.7 Conclusion	116
Chapter 5.....	117
Regional Policy & Impact Evaluation.....	117
5.1 Introduction	117
5.1.1 Theoretical background	117
5.1.2 Scope: Regional policy in England.....	117
5.2 Evidence types and the role of research credibility.....	118
5.3 The way in which evidence is incorporated into the policy making process.....	124
5.3.1 The RDA and Central Government evaluation model	124
5.3.2 Study designs and methods for RDA economic evaluation	129
5.4 Other factors besides evidence which affect the way policy is made	136
5.5 The RDA national impact evaluation.....	137

5.6 Conclusion	138
Chapter 6.....	140
Extending an EBM Approach to Regional Policy Evaluation	140
6.1 Introduction	140
6.1.1 Theoretical background	140
6.2 A comparison of EBM and regional policy evaluation	141
6.3 Epistemological challenges	146
6.4 Applicability challenges.....	149
6.4.1 Experimental methods.....	149
6.5 Recent developments for regional policy evaluation in 2015	154
6.4.2 Quasi-experimental methods	154
6.4.3 NICE for social policy	156
6.6 Conclusion	159
Chapter 7.....	161
The RDA Organisations: An Example	161
7. 1 Introduction.....	161
7.2 The principles of RDA policy making.....	161
7.2.1 The assessment of RDA success.....	161
7.3 The case study of Yorkshire Forward (YF).....	163
7.4 Yorkshire Forward: an example of EBPM?.....	163
7.5 Policy making processes in Yorkshire Forward	164
7.5.1 Organisational structure	164
7.5.2 Strategy	165
7.5.3 Project appraisal.....	166
7.5.4 Project delivery and monitoring	166
7.6 The generation of evaluation evidence in Yorkshire Forward.....	167
7.7 The communication and uptake of YF evaluation evidence.....	169

7.8 Conclusion	170
Chapter 8.....	172
The RDA Evaluation Experience: A Case of Evidence Based Regional Policy Making? .	172
8.1 Introduction	172
8.2 Theoretical background	172
8.3 Who were the survey respondents?.....	174
8.4 Presentation of the findings.....	174
8.5 The principles of regional policy making and the role of RDA evaluation.....	174
8.6 The uptake and use of evaluation evidence within the RDAs	176
8.6.1 Strategic decision making	176
8.6.2 Project appraisal.....	179
8.6.3 Delivery	182
8.7 The generation and communication of RDA evaluation evidence	184
8.7.1 The generation of RDA evaluation and the role of 'knowledge brokers'	184
8.7.2 The communication of RDA evaluation	186
8.8 Barriers and facilitators.....	189
8.8.1 A regulatory framework.....	189
8.8.2 Political backing.....	191
8.8.3 Analytical skill	193
8.8.4 Data quality and availability.....	194
8.9 A potential role for knowledge translation tools.....	200
8.10 Conclusion	201
Chapter 9.....	203
Using a Knowledge Tool to Extend an EBM Approach to Regional Policy Investment	
Prioritisation: A Critical Analysis	203
9.1 Introduction	203
9.2 Theoretical background	204

9.3 Who were the workshop participants?	205
9.4 Presentation of the findings.....	206
9.5 Investment prioritisation without the use of a decision support tool (scenario 1)	207
9.5.1 Strategic investment prioritisation in the RDAs.....	207
9.5.2 A hypothetical investment decision making scenario	208
9.5.3 An identified role for knowledge translation tools.....	212
9.6 Using a decision tool to support an EBRPM approach (scenario 2)	214
9.6.1 Receptivity to the decision support tool for investment prioritisation.....	215
9.6.2 The role of a decision support tool to shape policy decisions.....	217
9.7 The construction of a decision support tool	218
9.7.1 Drawing upon a knowledge base	219
9.7.2 Employing a decision logic	225
9.8 Developing the decision support tool.....	227
9.8.1 Developing the knowledge base: tacit knowledge	227
9.8.2 Multi Criteria Analysis	229
9.8.3 Updating data over time	231
9.8.4 Programmer skill and the need for capacity building.....	232
9.8.5 The role of a decision support tool to shape the generation of evidence ..	233
9.9 Conclusion	234
Chapter 10.....	237
Conclusion	237
10.1 Introduction	237
10.2 The research problem	237
10.3 The research questions and key findings.....	239
10.3.1 EBRPM and the role of evaluation	239
10.3.2 The role of knowledge translation tools and decision support	244

10.3.3 Research management insights from conducting the online workshop ...	246
10.4 Original contributions to knowledge	247
10.5 Limitations of this research.....	248
10.6 Implications for policy and practice	250
10.7 Recommendations for future research.....	251
10.8 Concluding remarks	252
Bibliography	253
WORD COUNT: 80,560 (89,047 including the Bibliography)	281
Appendix	1
Appendix 1: Example of RDA programmes, projects and links to IEF sub-themes	1
Appendix 2: Survey items and respondents - online survey.....	3
Appendix 3: Online survey respondents by RDA	25
Appendix 4: Example of RDA staff numbers by department.....	26
Appendix 5: Quantitative survey results.....	27
Appendix 6: Qualitative survey data tables	40
Appendix 7: Survey items and respondents - online workshop	68
Appendix 8: Additional screenshots of the online workshop.....	75
Appendix 9: Information provided for Scenario 1	77
Appendix 10: Follow-up telephone interview topic guide.....	80
Appendix 11: Online workshop participant information sheet	81
Appendix 12: Online workshop email communication.....	85
Appendix 13: Workshop participants	90
Appendix 14: Quantitative workshop results	91
Appendix 15: Qualitative workshop data tables.....	103

Figures

Figure 1: Coefficient of variation of GVA per capita in the English regions.....	4
Figure 2: An example policy cycle	36
Figure 3: Knowledge to Action Framework	43
Figure 4: Screenshot of the decision support tool presented in the workshop (user inputs and results).....	68
Figure 5: Screenshot of the decision support tool presented in the workshop (data background)	71
Figure 6: Screenshot of the decision support tool presented in the workshop (look-up tables).....	74
Figure 7: Screenshot of the workshop homepage	83
Figure 8: Screenshot of the workshop scenario 1 (top of the screen)	85
Figure 9: Screenshot of the workshop scenario 1 (bottom of the screen).....	86
Figure 10: Screenshot of the workshop scenario 2 (top of the screen)	88
Figure 11: Screenshot of the workshop scenario 2 (bottom of the screen).....	89
Figure 12: The Evidence Based Medicine Triad	113
Figure 13: MEANS classification of the types of evaluation (1995).....	119
Figure 14: IEF classification of the types of evaluation (2006)	120
Figure 15: RDA evaluation framework and methods of evaluative data collection.....	122
Figure 16: The Tavistock Guide classification of the types of evaluation.....	123
Figure 17: ROAMEF cycle	126
Figure 18: SQW analysis of evaluation methods and coverage.....	128
Figure 19: Distribution of votes	210
Figure 20: The context of the decision making process	213

Tables

Table 1: Criteria for judging quantitative and qualitative research.....	29
Table 2: Assessment criteria for the design of economic policy evaluation	31
Table 3: How will policy makers use evidence?	35
Table 4: Components of policy processes and different evidence issues	37
Table 5: Methodological framework.....	49
Table 6: Sampling framework	52
Table 7: Demographics of respondents	60
Table 8: Administering the workshop.....	97
Table 9: A hierarchy of strength of evidence for prevention and treatment decisions	103
Table 10: Examples of Outputs and Outcomes from the Treasury	131
Table 11: A comparison of EBPM debates across EBM and regional policy evaluation	142
Table 12: An epistemological comparison of EBM and regional policy evaluation	147
Table 13: A methodological comparison of EBM and regional policy evaluation	150
Table 14: Tasks for the What Works Centre for Local Economic Growth	157
Table 15: The Scientific Maryland Scale.....	158
Table 16: Demographics of online workshop respondents	206
Table 17: Example information presented to participants for scenario 1.....	209

Chapter 1

Introduction

1.1 Introduction

Have courage, citizens! We must go forward. But what are we aiming at? At government by knowledge, with the nature of things the only social force...

(Hugo 1862, p1004)

This quotation from Victor Hugo's acclaimed novel '*Les Misérables*' may be interpreted today as signifying a call for Evidence Based Policy Making (EBPM) built upon the foundations of scientifically rigorous research. In England, the concept of EBPM has been mobilised by the New Labour and Coalition governments through the 'what works' agenda, both *symbolically* as a means to legitimise and sustain political ideology with evidence portrayed in political discourse as apolitical, neutral and objective and *instrumentally* through attempts to embed evidence in policy making institutions and practices. Current debates about EBPM, and in turn debates about the role of evaluation evidence, are situated in a very particular historical, cultural and institutional context reflecting turbulent political and economic circumstances. A significant contention underlying these debates, which has certainly come to the fore recently, has focused on trust in the reliability of research findings leading to a call for the extension of approaches more closely aligned with Evidence Based Medicine (EBM) to wider social policy. Caution is needed, however, as EBPM is a contested concept across policy domains, and within social science (Wells 2007). There is a lack of consensus in academic and policy literature on the role of evidence in policy making with debates centring on: what kinds of evidence are used and the role of research credibility; the issues surrounding the way in which evidence is incorporated into the policy making process; and what are the other factors besides evidence which affect the way policy is made.

This thesis tests the applicability of, and the factors which determine the success of, Evidence Based *Regional* Policy Making in England. The regional policy context

provides a fascinating, and previously under-researched case study to explore the wider EBPM debates given the complexities associated with its multifaceted policy agenda, structures and mechanisms, alongside its inherently political character.

This first chapter is an introduction to the problem studied. It begins with background information on the context for the study and the rationale for place-based policy and the evolution of regional policy in England influencing the current situation for local growth policy evaluation. An overview of the conceptual framework is then presented, describing the relationship of the research to existing literature and theories. Next the research problem is defined through highlighting gaps in the current knowledge base and stating how the research aims to address these. Two major aims are stated and these aims are expressed in terms of three research questions. The need to adapt to the evolving study context and my place vis-à-vis the research is then reflected upon. Finally, an overview of the structure of the thesis is given, leading the way into the succeeding chapters.

1.2 Context: Regional policy in England

It must be acknowledged at the start of this thesis that English¹ regional policy is highly contested and controversial, which makes it an appropriate topic for scholarly inquiry and debate. Since 1928, a widely accepted justification for regional policy has been set out as follows: regional economic disparities (e.g. in unemployment rates, per capita incomes and living standards), which persist for long periods of time, have harmful effects on the national economy and may have harmful political and social consequences. Therefore public policy is required to address these disparities (Armstrong and Taylor 2000). However, it could be argued that whether or not policy makers (and academics) judge that there is a sound case to invest public funds into regional policy, or some other type of spatial policy, and the form that such policy instruments should take, is inherently *ideological*.

There is considerable disagreement concerning: the multitude of analyses and theoretical frameworks on regional development to inform the rationale (or not) for intervention; the balance between potentially conflicting objectives for rebalancing the

¹ The role of evaluation evidence within the *English* Regional Development Agencies is the main focal point of this research. However, literature discussing British, UK and English regional policy will be discussed.

economy and sustaining national competitiveness² and between economic and social objectives; and the role of the state and the institutions and interventions that are needed to achieve such objectives. Barca (2011) explains that such differences in regional development thinking reflect fundamentally different philosophical understandings of the economy, the state and the wider community and the relationships between them. It is not within the scope of this thesis to summarise the whole breadth of thinking across these issues, or indeed to set out a case for regional policy, however, the key aspects of regional policy development are considered below to give necessary context to the subsequent discussion on regional policy evaluation.

1.2.1 Spatially unbalanced growth in England

The main function of regional policy in England has traditionally been remedial or curative, to address spatially unbalanced growth. Key studies exploring the long-term trends of spatial economic growth and competitiveness in England have tended to focus on the spatial level of the region and have employed methods to empirically analyse whether or not regional imbalances are increasing, termed 'divergence', or reducing, termed 'convergence' by calculating coefficients of variation. Crafts (2005, p61) estimated historical regional GDP³ per capita in Britain and revealed that London's per capita GVA⁴ has been consistently higher than in the rest of the country over the last 140 years (at least), that after the First World War to the 1970s there was a sustained episode of regional convergence and that between 1970 to 2001 there was a period of rapidly rising regional imbalances. Recent government analyses of the ONS⁵ Regional Accounts indicate that this trend of divergence has continued (BIS 2010b, p37). Other studies comparing regional convergence rates across industrialised countries, via cross-sectional regression analysis, have generally supported Crafts' findings for the UK (Barro and Sala-i-Martin 1991; BIS 2010b, p38).⁶

² Garretsen et al. (2013) demonstrate, however, that evidence for such a policy trade-off is ambiguous.

³ Gross Domestic Product.

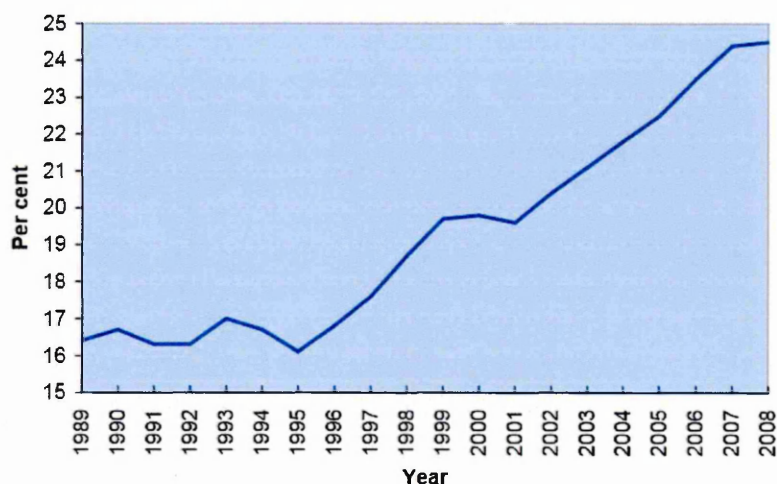
⁴ Gross Value Added.

⁵ Office for National Statistics (ONS).

⁶ The comparative findings of such studies have been ambiguous. Recent government analysis (BIS 2010b) updated Barro and Sala-i-Martin's analyses using OECD Regional Accounts Data. According to this analysis, all five industrial countries considered (UK, US, Germany, Italy and France) seem to have seen at least a slowdown in the rate of

Regional data indicate that London's economic performance significantly outstrips the other regions of the UK (BIS 2010b), as shown in Figure 1. "In 1989, dispersion between the regions (as measured by a coefficient of variation) was around 16 per cent but by 2008 this had increased to over 24 per cent" (BIS 2010b, p7). However, adding to the complexity, Garretsen et al. (2013, p180) argue that economic growth rates cannot simply be explained by a "core-periphery story."⁷

Figure 1: Coefficient of variation of GVA per capita in the English regions⁸



Source: BIS 2010b, p7

What is to be done about such regional disparities, and the continuing identified trend of divergence, depends on whether or not government policy making can have a bearing on the factors which influence local economic growth. Crafts (2005) identified that the long-run trends in England are entrenched in deep historical roots originating back to the industrial revolution and the resulting urban settlement patterns of former industrial or manufacturing centres. Therefore, regional imbalances are identified to be strongly influenced by globalisation (Crafts 2005) and exposure to improving

convergence and a reversal in the UK and the US since the earlier period, 1950–1985, studied by Barro and Sala-i-Martin (1991). The BIS paper (2010b) further reports that Germany and Italy experienced a decrease in regional imbalances, France experienced neither increasing nor decreasing imbalances, while the United Kingdom and the United States experienced increasing imbalances for the period 1995–2007.

⁷ Garretsen et al. (2013, p180) draw attention to evidence which suggests that some non-core regions have begun "to account for an increasing share of economic growth across many OECD countries (and most markedly in Europe)."

⁸ Calculations using ONS Regional Accounts for English NUTS1 areas 1989–2008.

technology (BIS 2010b). Crafts (2005) explains that this has promoted deindustrialisation in the Midlands and North of England, leading to a fall in demand for unskilled labour, while favouring the growth of business and financial services in London and the South East, leading to growth in demand for skilled labour. Such an interpretation is supported by others, and the associated roles of trade liberalisation (Collier and Dollar 2001) and the current economic crisis (Gardiner, Martin and Tyler 2012) have been identified as key factors.

It could be argued, then, that any government's policy to address regional disparities is heavily impacted upon by the wider context of the dynamic world economy. Building understanding of the underlying drivers of regional growth and spatial disparities, which government policy making can aim to influence, has therefore been a prevailing subject of academic consideration. Such evidence has been used to inform, as well as legitimise, government spatial policy.

1.3 A brief history of regional policy in England: 1928 to the present day

1.3.1 Aims

Historically the aims of regional policy have tended to change periodically within and between government administrations leading to policy switches. Armstrong and Taylor (2000) undertook a detailed analysis of the chronological progression of regional policy since its inception in 1928 up to the end of the 1990s and identified the main 'phases' and key features (p214-225) as well as the characteristics of the 'free market' versus the 'interventionist' approach to regional policy making (p210-213). The timeframe is partially overlapped and extended beyond Armstrong and Taylor's work in contemporary analysis by Kitson (2012) and in a paper by Grimshaw and Rubery (2012) studying the UK social model under both New Labour and the Coalition. Such enquiry is briefly summarised and expanded upon in the following paragraphs to provide background contextualising information.

Regional policy began during the inter-war years in Britain in response to the depression. The objective of reducing regional disparities in unemployment became the guiding principle underlying subsequent policy going forward (Armstrong and Taylor 2000). A *Keynesian*, interventionist 'welfare state' model of government was developed post-war (Clarke 1988) and the White Paper 'Employment Policy' (1944)

made a commitment to the attainment of full employment. At the centre of the interventionist approach was the view that the 'regional problem' was caused by structural weakness in the regional economy (Armstrong and Taylor 2000) alongside a drain of financial capital from poor to rich regions (Martin and Minns 1995). Emphasis was placed on the need to create jobs in areas of traditionally high unemployment, based upon the findings of the Barlow report (1940). Development Areas were established and policy instruments were introduced such as loans and grants to firms, and placing controls on the location of industry (Armstrong and Taylor 2000). Regional policy waxed and waned over the coming decades, but the main thrust remained broadly the same. Notably, in the 1960s two long-standing concerns for government policy began to emerge: the rate of national growth compared with other industrialised countries; and the impact of excessive growth in Greater London (Armstrong and Taylor 2000).

At the same time as these policy developments, early neoclassical theories of growth began to emerge in the academic literature (Solow 1956)⁹ predicting economic convergence across regions, as long as economic markets were functioning well and resources and technology were mobile. The neoclassical approach provided a rationale for tackling market failures.

The election of the Thatcher government in 1979 forced a memorable U-turn away from Keynesian demand management strategies towards monetarism and a *neoliberal* model of government, rolling back state intervention and placing emphasis on the free market economy (Clarke 1988). Despite cuts in the levels of support, regional policy survived this turbulent period and there was a shift of focus towards selective assistance and the encouragement of enterprise (Armstrong and Taylor 2000). In the late 1980s and early 1990s, 'New Growth' theories began to develop in the academic

⁹ Early (exogenous) neoclassical models theorise that "people will move to areas with high capital/high productivity to receive higher wages; firms on the other hand will move to low capital/low productivity areas to receive a higher return on their capital investment. Such movement will continue until workers and investors respectively receive a similar return irrespective of their location. That is, the spatial equilibrium would occur when all areas converged to a similar level of productivity" (BIS 2010b, p19-20). The model assumes complete factor mobility, including the diffusion of technological advances.

literature (Romer 1986; Lucas 1988; Rebelo 1991).¹⁰ These theories acknowledged 'endogenous' factors driving regional development and suggested that long-run growth emanates from investment in human capital which has spillover effects on the economy (Gardiner, Martin and Tyler 2004; 2012). In 1988, the role of regional policy was radically changed towards the objective of 'indigenous development' (i.e. self-sufficient growth) through the removal of 'supply side' economic rigidities, heavily influenced by EU regional policy (Armstrong and Taylor 2000). By the 1990s regional policy had become firmly entrenched within Britain's national industrial policy with a focus on enhancing national competitiveness (Armstrong and Taylor 2000).

The election of New Labour in 1997 signified a shift back towards an 'interventionist' model of government. However, instead of taking a Keynesian demand management approach (Kitson 2004), efforts were directed to the '*supply side*' and the rebuilding of industrial and commercial bases in 'problem regions' (Armstrong and Taylor 2000). A strong commitment was made towards maintaining competitiveness (HM Treasury 2003a) and Gordon Brown (cited by White 1994) famously described the economic approach as rooted in ideas of "post neo-classical endogenous growth theory." New Labour identified skills, enterprise, innovation, competition and investment as 'five drivers' of productivity (HM Treasury 2001; 2003a). During the late 1990s, however, there was a slow shift in the focus of regional policy towards social objectives with emphasis placed on the concept of "social exclusion" (Armstrong and Taylor 2000, p226). In 1999, the Cabinet Office published a paper titled 'Sharing the Nation's Prosperity' which provided evidence of the multifaceted nature of spatial imbalances, going beyond purely economic indicators to include measures such as indices of multiple deprivation and educational attainment. However, critics highlighted the deep-seated tension between social and economic rationales for intervention (Grimshaw and Rubery 2012, p106). The result was that regional policy did not have one main objective (e.g. job creation, GDP increase), but "multiple, potentially

¹⁰ Gardiner, Martin and Tyler (2004; 2012) provide comparison to earlier models. "In such theories, there is no prediction that economies with different performance levels are likely to converge" (BIS 2010b, p20). In these theories spillover effects reduce the diminishing returns to capital accumulation.

conflicting” social, economic and environmental objectives and various streams of intervention (Armstrong and Taylor 2000, p231).

The influence of EU regional policy

In 1973 Britain became a full member of the European Economic Community. This marked a turning point for regional policy and enabled the assisted areas to benefit from various expenditure streams including European Social Fund (ESF) grants (Armstrong and Taylor 2000). A significant event was the establishment of the European Regional Development Fund (ERDF) in 1975 (part of the agreement of the UK's accession) which was designed to complement, rather than replace national regional policy (DG REGIO 2014a). Later, and in response to the strains imposed on disadvantaged regions by the twin processes of widening and deepening the Single European Market, new EU legislation reformed all aspects of regional policy between 1988 and 1993 (Armstrong and Taylor 2000). In 1988, the new '*Structural Funds*' brought together the ERDF, ESF and the EU's Agriculture Funds. Armstrong and Taylor (2000) reflect that this signified a switch from piecemeal project-by-project financing to the strategic orientation of investments and the implementation of co-ordinated multi-annual programmes. In 1993 the Maastricht Treaty introduced the Cohesion Fund and between 1994 and 1999 the resources for the Structural and Cohesion Funds (Cohesion Policy) were doubled, to equal a third of the EU budget (DG REGIO 2014a). Expenditure on EU regional policy has continued to grow, accounting for €347bn, or 35.7% of the total EU budget for 2007-13 (DG REGIO 2014a).

There has been a complex relationship between national and EU regional policy. EU regulation states that member states must provide match funding to draw down EU funding, ensuring a continued role for national governments. In addition to match funding EU programmes, the UK has continued to have its own *national policy* instruments. However, Armstrong and Taylor (2000) reflect that the outcome of the reforms to EU regional policy has been to place the EU in the 'driving seat'. During the 1990s, the aims of EU regional programmes tended to be predominantly in harmony with British regional policy, focusing towards indigenous development and changing the supply side of the economy, with a strong focus on innovation policy (Morgan 1997). Despite this, ongoing tensions are possible given that the UK government

continues to use national (domestic) regional policy instruments to attract inward investment from overseas, whilst EU competition policy sets out strong injunctions against unfair competition and regulates the role of State Aid to prevent “subsidy wars” (HM Government 2014, p42).

In 2000, “the ‘Lisbon Strategy’ shifted the EU's priorities towards growth, jobs and innovation and the priorities of cohesion policy were shifted to reflect this” (DG REGIO 2014a) (see Mendez 2011). Of the €347bn budget for the programming period 2007-13, 25 per cent was earmarked for research and innovation, and 30 per cent for environmental infrastructure and measures to combat climate change (DG REGIO 2014a). However, currently there is an uncertain future for the financing of EU cohesion policy for the programming period 2014-20 due to macroeconomic conditions and potential co-financing constraints across Europe. Begg et al. (2014, p16) have recently reported that “although structural reforms have notionally been centre-stage since the launch of the Lisbon Strategy in 2000, with the bulk of the Europe 2020 strategy being about changing the supply side of the economy, they have become less prominent in the policy discourse in the last two years.” Recent debates have arisen around a potential shift in the paradigm of EU regional development thinking (e.g. Barca 2009). Garretsen et al. (2013, p182) explain that the latest EU regional policy framework embraces a “modern plea for place-based regional policy that tries to strengthen the endogenous growth potential of lagging and peripheral regions.”

Current situation in 2015

Following the formation of the Coalition government in 2010, regional policy in England was completely overhauled with emphasis firmly switched back to issues of national competitiveness over equity (BIS 2010b). Grimshaw and Rubery (2012) contend that the Coalition has sought to embed “a stronger neoliberal approach to social policy” (p105), characterised by the withdrawal of the state towards a “liberal market economy with a *residual welfare state*” (p107). Taylor-Gooby and Stoker (2011, p14) concur, adding that the shift amounts to more than ‘politics as normal’, arguing that it involves “rolling back the state to a level of intervention below that in the United States - something which is unprecedented.” A programme of austerity policy reforms have been implemented, largely impacting on public sector jobs

(Grimshaw and Rubery 2012). The concept of 'rebalancing' the economy has shifted in policy discourse from rebalancing across the regions towards sectoral and public/private rebalancing as well as spatial rebalancing (Gardiner, Martin and Tyler 2012, p4). The region has been denounced as an arbitrary administrative boundary and a place-based approach has been taken, focused on 'functional economic geographies' (BIS 2010b). The collection of regional statistics has been stopped (Ferry and Bachtler 2013) and some scholars have argued that the word 'region' is being removed from the contemporary English policy vocabulary (Bentley, Bailey and Shutt 2010; Pugalis 2011).

The Cabinet Office's evidence paper 'Understanding Local Growth' (BIS 2010b, p23) suggests that the Coalition's world view and the spatial scales favoured for policy have been influenced by New Economic Geography (NEG) theories (Krugman 1991),¹¹ drawing upon micro level analysis of the spatial economy (Venables 2008) and the concept of 'agglomeration economies'.¹² Interestingly, the evidence paper states that "even with fully functioning markets, there can be an uneven distribution of economic performance and persistent differences that are not necessarily due to market failure" (BIS 2010b, p23). Later in the paper it is argued that there "may be substantial limits to how geographically balanced an economy may become" (BIS 2010b, p26). Such a challenge to the very basis of a place-based approach to policy (let alone regional policy) has also been echoed in the academic literature, most notably by Overman (2013), the director of the 'What Works Centre for Local Economic Growth'. Overman and Gibbons (2011, p24) have contended that "disparities are driven by people rather than place." Garretsen et al. (2013, p181) argue that focusing regional policy on encouraging people (and firms) "to migrate to and succeed in economic centres" possibly only strengthens core-periphery patterns.

¹¹ The clustering of economic activity, generating an uneven distribution of activity and income across space, is a prevalent finding of this stream of research (Krugman 1991). Many of the NEG models predict increasing regional specialisation as both people and firms move to areas of high productivity (Gardiner, Martin and Tyler 2004).

¹² Agglomeration theory suggests that concentrations of economic activity generate economic benefits for the firms located within them including: a supply of labour on which firms are able to draw (Glaeser and Resseger 2010); easier access to inputs and suppliers (Puga 2010); and the creation of knowledge spillovers (Audretsch and Feldman 1996; Jaffe, Trajtenberg and Henderson 1993).

1.3.2 Institutions and participants

As highlighted above, regional policy has been exposed to short-term political pressures. This has led to institutional churn and, more recently, 'institutional termination' (Ferry and Bachtler 2013). However, for a long part of its history regional policy in Britain was the virtual monopoly of the national government (Armstrong and Taylor 2000). This monopoly effectively ended after the establishment of the ERDF in 1975 and was cast asunder by EU reforms in 1988 when the EU Structural Funds became the key driver of UK regional policy. The 1988 EU reforms committed to a partnership approach to regional policy and the 1993 Maastricht Treaty entrenched the principle of subsidiarity.¹³ New Labour's 'reconstructed Keynesianism' approach also necessitated multi-level involvement at the local, regional, national and EU levels of government (Armstrong and Taylor 2000). Although the foundations for a regional institutional framework in Britain were established by the Major government, with the formation of the regional Government Offices in 1994 (Bache 1998), Lloyd and Meegan (1996, p75) contend that Government Offices "maintained a strong bias toward central government control at every stage." Hayward (1997, p378) notes that there was a "highly discredited challenge culture that... led to regions being forced to engage in open competition with other regions to receive their share of the national pot of money for economic development."

Regional Development Agencies

New Labour bolstered regional decision making with increased administrative responsibilities given to regional Government Offices and later establishing *Regional Development Agencies* (RDAs) for strategic planning and economic development (Great Britain 1998). The Labour Party Manifesto (1997) set out an overarching vision for the RDAs to "co-ordinate regional economic development." Although the RDAs were financed by national government public funds¹⁴ via the creation of a 'single pot'

¹³ The subsidiarity principle aims to ensure that decisions are taken as closely as possible to the citizen (DG REGIO 2014a).

¹⁴ In addition to European Regional Development Funding.

of RDA funding,¹⁵ there was a high degree of flexibility, enabling funding to be targeted towards the investments the RDAs prioritised for each region (YF 2009).

“Accountability, effectiveness and subsidiarity” were set out as clear policy making principles in the 1996 ‘Report of the Regional Policy Commission’ (cited by Hayward 1997, p378). Regionally, the RDAs were initially accountable to indirectly elected Regional Chambers made up of regional partners including “representatives of local authorities, economic and social partners (e.g. business associations, trade unions and voluntary groups) and other sectoral interests (e.g. higher education, environment and rural)” (Pike et al. 2012, p104). RDAs were also directly accountable for the way in which they used their resources nationally through their sponsor department (BIS) and for delivering effectively against Regional Economic Performance (REP) Public Service Agreement (PSA) monitoring targets set by central government. In particular, PSA7 tasked the RDAs with improving the economic performance of all English regions and reducing the gap in economic growth rates between regions (HM Treasury 2003a). The most recent reporting regime monitored performance in terms of ONS Regional Accounts GVA estimates (ONS 2011).

At an operational level, delivery of PSA7 was implemented via Regional Economic Strategies (RES). RDA’s were also considered an appropriate institutional framework to operate ‘indigenous development’ policies and thus their objectives were wide-ranging: “furthering economic development and regeneration; promoting business efficiency and competitiveness; promoting employment; enhancing the development and application of skills relevant to employment; and contributing to sustainable development” (Great Britain 1998, p8).¹⁶ Often the RDAs sought to achieve their objectives via funding projects through local level ‘delivery’ organisations, as a means of enabling the active participation of the local community. Polverari and Bachtler

¹⁵ ‘Single Pot’ pooled money from all the contributing government departments in the UK: The Department for Business, Innovation & Skills (BIS); The Department for Communities and Local Government (CLG); The Department for Innovation, Universities and Skills (DIUS); The Department for Environment, Food and Rural Affairs (Defra); The Department for Culture, Media & Sport (DCMS); UK Trade & Investment (UKTI). BIS was the sponsoring department (YF 2009).

¹⁶ See Appendix 1 for an example of RDA programmes and projects.

(2004, p12) note that the number of actors and mechanisms involved in policy making became “unprecedented in comparison with the past.”

In 2004 plans to enhance the accountability of regional institutions, through an elected regional assembly, were rejected and subsequently the ‘Sub-National Review of Economic Development and Regeneration’ (SNR) process was introduced in 2007 to streamline state involvement in regional policy (Pike et al. 2012). This led to the abolition of the Regional Chambers and placed emphasis upon sub- and city-regional partnerships and joint working, Regional Ministers, a parliamentary regional select committee (Pike et al. 2012) and the delivery of PSA7 through integrated strategies (HM Treasury 2007). Leading up to the UK General Election, Pike et al. (2012, p104) notes that “SNR collided with the assessment and emergent critique of New Labour’s approach.” The first RDA national evaluation was published (PWC 2009a) and, despite presenting broadly positive conclusions on the effectiveness and efficiency of the RDAs, Ferry and Bachtler (2013, p269) note that “these were ignored and indeed contradicted by the political narrative.”

Current situation in 2015

The formation of the Coalition government in May 2010 led to a dramatic and rapidly evolving change of policy direction and to austerity cuts.¹⁷ Ministers in the new Coalition government denounced the previous regional institutional framework as wasteful, bureaucratic, unnecessary and ineffective (Ferry and Bachtler 2013). There was radical transformation in the structures and funding mechanisms to support local growth; thirteen years of ‘experimentation’ with a regional tier was brought to an end with the abolition of the RDAs by March 2012 and other parts of regional policy administration were dismantled (Ward and Hardy 2012). The Coalition further devolved power to Local Enterprise Partnerships (LEPs) and introduced the ‘Regional’ Growth Fund (RGF) in 2010, and Growth Deals in 2014 as well as other initiatives to be discussed below. Alongside the rationale of an institutional shift from ‘regionalism’ to ‘localism’, such policy change was reasoned to be a response to cyclical budgetary constraints and introduced as part of a package of austerity measures. However, the

¹⁷ The Emergency Budget announced £270m cuts to be found by the RDA Network during 2010/11 (YF 2011a, p4).

initial scale of the cuts meant that the total level of funding available was approximately one third of the RDAs' budget (NAO 2013b, p9). Grimshaw and Rubery (2012, p121) contend that such "above-average cuts to local government revenues (27% over four years) started a process of downsizing workforces and likely long-term loss of competences in many activities, including... economic planning."

At the time of writing, thirty-nine LEPs have been established (BIS 2013). They cover areas intended to relate to functional economic market areas, however, criticism has been raised that LEP boundaries are still 'arbitrary' (Townsend 2012). In terms of funding, the RGF has been set up to influence private sector employment and to lever in private sector investment. Economic appraisal for the RGF has been conducted through six competitive funding rounds. For the fifth and sixth rounds, there has been a change in the emphasis of the fund, with eligibility limited to applications led by the private sector only (i.e. LEPs were not able to bid). When these final two rounds close, the total funding commitment to RGF is expected to be £3.2bn up to March 2017 (BIS 2014a). However, RGF coverage of the country has not been universal, and to meet the need for LEP funding, Growth Deals have recently been introduced and were 'signed off' in July 2014 on the basis of agreed strategic economic plans. However, it could be argued that the Coalition has been deliberately vague about the amount to be invested via Growth Deals, and how much 'new money' this constitutes.

It remains to be seen whether or not the new institutions and policy instruments intended to influence local economic growth will be successful. In 2013, the LSE Growth Commission proposed the creation of an "independent National Growth Council to review relevant evidence and to recommend growth-enhancing policy reforms" (LSE 2013, p2). The resulting report (LSE 2013, p34) signified a call for EBPM and policy evaluation:¹⁸

We must break the familiar cycle of institutional churn and political procrastination and find ways of ensuring that difficult and contentious long-term decisions are based on the best available independent expertise.

¹⁸ Likewise for EU regional policy, Bachtler, Méndez and Vironen (2014, p52) note that "in the context of the ongoing criticism about the effectiveness of Cohesion Policy, a key challenge is to ensure that it produces quantifiable results and impacts and that it visibly and measurably contributes to the Europe 2020 strategy."

Overall, the shift in regional policy thinking from a Keynesian model to one which is focused on issues of competitiveness, local institutions and place-based factors means that evaluation of regional policy is at somewhat of a crossroads. In addition, scholars have raised the question as to whether the weakened evaluation culture will survive the reduction or withdrawal of Structural Fund intervention (McNamara et al. 2009).

1.4 Regional policy evaluation in England: the current situation in 2015

It has been widely accepted for a long time that evaluation is essential if regional policy is to be efficient, effective and meet its objectives (Armstrong and Taylor 2000). This feeds into a wider debate about the role of evidence in policy making. In England, the concept of EBPM has been mobilised by the New Labour and Coalition governments through the 'what works' agenda. EBPM has gained political currency since 1997 when the incoming Blair government claimed that policy was to be shaped by evidence and rational decision making, implying that the "era of ideologically driven politics was over" (Nutley 2003, p3). The mantra '*what counts is what works*' was declared in New Labour's party manifesto for the 1997 General Election (Labour Party 1997) and a bold commitment was made in the 1999 White Paper 'Modernising Government' that "policy decisions should be based on sound evidence" (Cabinet Office 1999a, p31). The role of research and evaluation were underlined:

Good government is thinking government... rational thought is impossible without good evidence.... social science research is central to the development and evaluation of policy.

(David Blunkett, UK Minister for Education, cited in Nutley 2003, p3)

'Professional Policy Making for the Twenty First Century' (Cabinet Office 1999b) set out a commitment to: using the best available evidence, building evaluation into the policy process and learning from experience of 'what works' and 'what does not work'.

However, Wells (2007) identifies that over time New Labour's attitude to EBPM was shaped by a shift from a focus of policy learning and experimentation towards policy delivery, and thus greater attention was placed on 'hard' quantitative and economic analysis. Perhaps the pinnacle of the EBPM pursuit was the establishment of the National Institute for Clinical Excellence (NICE) for health policy, tasked with systematically appraising evidence on cost effectiveness, alongside the large-scale

national evaluations of Business Link (Mole et al. 2008), the New Deal for Communities Programme (CLG 2010a; 2010b), and Sure Start (Belsky, Barnes and Melhuish 2007).

In terms of regional policy evaluation, the influence of the 'what works' agenda is examined in detail in Chapter 5. Suffice to say that despite emphasis being placed on evaluation, a report by the NAO¹⁹ (2010, p7) declared that RDA evaluation was "weak" and remarked that: "we are unable to conclude that the regional wealth benefits actually generated were as much as they could and should have been, and are therefore value for money." This has led to a call for lessons to be learnt from the RDA evaluation experience (Great Britain, Parliament, House of Commons 2010; Chadwick, Tyler and Warnock 2013; Garretsen et al. 2013). Despite this, the Coalition's new local growth funds and structures have not been designed as a co-ordinated national programme and initially the NAO report 'Funding and structures for local economic growth' (2013b, p11) challenged that:

The government does not have a clear plan to measure outcomes and evaluate performance and therefore show value for money across the programme. As a result, departments cannot be sure about where to direct their resources to achieve the most impact. Although individual initiatives monitor their progress it is not done in the same way across initiatives. Consequently, this does not present an overall comparable picture of performance.

The current situation for regional policy evaluation at the time of writing is captured within the BIS Evaluation Strategy for 2015-16 (BIS 2014b). A scoping study to develop a framework for evaluation has recently been completed for the RGF, the full evaluation has been commissioned and is currently underway, and a scoping study has recently been commissioned for the Growth Deals evaluation (BIS 2014b, p16-17). Of interest is that the Growth Deals evaluation aims to explore options for potential "cross-cutting" evaluation, inferred to be *across* LEP areas (BIS 2014b, p17). Overall, the strategy calls for the use of innovative methodological approaches to identify additional local economic growth and in particular states that "[BIS] look for opportunities where impact evaluation techniques such as Randomised Control Trials (RCT) or quasi-experimental designs can be used" (BIS 2014b, p6). This pursuit of identifying credible and robust counterfactuals is evident with a "matched with before and after" study design proposed for the RGF, alongside econometric analysis and the

¹⁹ National Audit Office

use of economic evaluation (i.e. the monetisation of costs and benefits and the reporting of a cost-benefit ratio) (BIS 2014b, p16).

The outcomes of such evaluation planning remain to be seen, however, a significant underlying debate has emerged focused on trust in the reliability of research findings. Indeed, the recent 'Evaluation in Government' report (NAO 2013a) was critical of the level of utilisation of evaluation evidence and was also critical of the historic reliance on 'lower-power' methodologies in the evaluation of business and spatial interventions.

It could be argued that this focus on trust in research findings has led to a call within recent policy discourse for the extension of approaches more closely aligned with *Evidence Based Medicine (EBM)* to wider social policy. The most obvious evidence for this is the Cabinet Office paper 'Test, Learn, Adapt' (Haynes et al. 2012).²⁰ This paper was famously called the "Ladybird Book of RCTs" (Goldacre 2013), and claims the use of Randomised Controlled Trials (most often applied within medical research) should be extended across government policy to "pinpoint cost-effectiveness" (Haynes et al. 2012, p12). A variant of the 'trust' argument has been that "single-study findings are misleading, and that a better understanding of causes and consequences emerges from systematic reviews of all available research" (Head 2008, p17). Thus the Coalition's establishment of the 'What Works Centres' in 2013, conceptualised collectively as a "NICE for social policy" in policy discourse (Cabinet Office 2013, p1), has also been inspired by the institutional framework for EBM policy making within the healthcare sector and asserts a similar focus on high quality impact/economic evaluation evidence. Caution is needed, however, as EBPM is a contested concept and Evidence Based *Regional* Policy Making (EBRPM) in England is an under-researched area. Thus the implications of extending such an EBM approach have been insufficiently examined given the recent shift of focus for regional policy evaluation within the wider 'what works' agenda.

²⁰ Also see Torgerson and Torgerson (2008).

1.5 Defining the research problem

1.5.1 The conceptual framework: Evidence Based Policy Making

The call for EBPM “reflects an ambition to deliver better policy in terms of outcomes, resource efficiency and effectiveness, and a belief that this can be achieved through utilising the available evidence to inform and guide decision making” (Rutter, Hawkins and Parkhurst 2013, p2). The traditional case for EBPM is set out in the literature as follows: that policy making and professional practice should not be ‘opinion based’ (Gray 1997), which Davies (2004, p3) notes “relies heavily on either the selective use of evidence (e.g. on single studies irrespective of quality) or on the untested views of individuals or groups, often inspired by ideological standpoints, prejudices, or speculative conjecture.” Instead, “the pursuit of EBP[M] is based on the premise that policy decisions should be better informed by available evidence and should include rational analysis” (Sutcliffe and Court 2005, p3). Stemming from the Evidence Based Medicine (EBM) movement, described as a “new paradigm for medical practice” (EBM Working Group 1992, p2421), overall a more rigorous and systematic approach is advocated. There is a wide body of literature focusing on the role of evidence in the policy making process; however, as Wells highlights (2007, p23) “there is no single unifying account of EBPM. It is used in different ways across the policy and academic worlds” (see also Davies, Nutley and Smith 2000).

By reviewing the EBPM literature (Chapter 2), two knowledge gaps are identified. Firstly, a dominant perspective within EBPM literature has been to use EBM approaches as a yardstick against which wider social policy is assessed and to draw parallels between the practices of EBM and EBPM (Sefton 2000; Sefton et al. 2002; Sefton 2003; Dobrow, Goel and Upshur 2004; Cookson 2005; Somekh et al. 2005). However, there has not been such a study exploring the extrapolation of an EBM approach to the regional policy context specifically. Secondly, a strand of the EBPM literature focuses on how to achieve the maximum impact for a body of evidence and on the strategies which may be employed to improve the uptake and use of research. However, the utilisation of knowledge translation tools and decision support tools as a mechanism to support the uptake of evidence is an emerging topic in the EBPM literature (Straus, Tetroe and Graham 2013).

To examine EBPM debates across policy domains, a conceptual framework was developed to structure the analysis. In particular, a paper by Sutcliffe and Court (2005) was drawn upon which distinguished between three central theoretical questions in the EBPM literature: what kinds of evidence are used and the role of research credibility; how is evidence incorporated into policy making; and what are the other factors besides evidence which affect the way policy is made. These three cross-cutting conceptual questions, in addition to the potential use of a knowledge tool to extend an EBM approach to regional policy (i.e. as a strategy to increase the uptake of evidence), are examined in turn throughout this thesis.

1.5.2 Gap in knowledge: Evidence Based Regional Policy Making

The regional policy context provides an excellent case study to explore the wider EBPM debates given the complexities associated with its multifaceted policy agenda, structures and mechanisms, alongside its inherently political character (section 1.2 and 1.3). Although regional policy has subsets and related fields (such as small business policy, skills policy and infrastructure planning), and a number of subnational institutions deliver economic policies, the scope of this research focuses on regional policy as delivered through the RDAs. A key issue that defines the research problem is that although the academic literature is well established and becoming increasingly sophisticated, the practice of regional policy evaluation in England by the institutions charged with its implementation has not kept pace with this development. Less attention has been paid to the evaluation of *UK regional policy instruments* (i.e. non EU policy instruments) and to the *processes* of undertaking regional policy evaluation in practice (to be discussed below). Thus, given the recent shift of focus for regional policy evaluation within the wider 'what works' agenda, the implications of extending an EBM approach are relatively unknown.

Although there has been a general trend of regional data and regional policy evaluation strengthening in both theory and method over time, the influence of EU regional policy and the subsequent demands to evaluate Structural Fund expenditure from 1989 means that EU economic agencies have been at the forefront of developing the methodology for evaluation practice (i.e. MEANS; Evalsed) and a common set of guidelines for the monitoring and evaluation of EU programmes within and across regions (Bachtler and Michie 1995). The EU has undertaken and published an extensive

body of ex-ante, ex-post, and more recently on-going and thematic evaluations of its programmes (DG REGIO 2014b). Regional policy evaluation in England has therefore inclined towards the European Structural Funds (e.g. Polverari and Bachtler 2004; Bachtler and Wren 2006;²¹ Bachtler 2011).

Certain key aspects of UK regional policy evaluation have tended to be overlooked. For instance, while decentralisation of funding and powers to the regional and then local levels has been a key focus of New Labour and then Coalition policy, a study has not yet been undertaken to assess whether such institutions and policies have (or will) achieve superior outcomes to a more centralised approach and to examine the national efficiency of regional policy. It could be argued this suggests a lack of acknowledgement that regional policy is a national policy competency. The case for strong central government control over regional policy has been examined in the literature, however. Armstrong and Taylor (2000, p342-343) put forward four key arguments: central government has a legitimate interest in seeking solutions to regional problems; central government control is needed to ensure that regional policy is adequately funded in the regions of greatest need; central government involvement is necessary to ensure the effective co-ordination of regional policy; and central provision of a regional policy instrument is sometimes more efficient.

In terms of understanding the processes of regional policy evaluation in England, and the realities, possibilities and challenges of evaluation evidence, the NAO had formally reviewed the evaluation functions of the RDAs via an 'Independent Supplementary Review' (ISR) process. However, further discussion is mainly presented in departmental (e.g. ONS 2011, NAO 2013a) or professional reports (Cook et al. 2008) rather than in the academic literature. Although there have been attempts to understand the utilisation and influence of evaluation within central (NAO 2010; Rutter 2012; BIS 2014b) and local government (Percy-Smith et al. 2002; Allen, Grace and Martin 2014), less attention has been paid to the processes of evaluation within the RDAs. To the researcher's knowledge, one recent paper by Chadwick, Tyler and Warnock (2013) stands alone in examining the processes of impact evaluation within the RDAs and

²¹ A special edition of the Regional Studies journal was focused on the evaluation of Cohesion Policy (Bachtler and Wren 2006).

describes lessons learnt for the evaluation of LEPs and the RGF in times of austerity.

The authors contend (2013, p844):

It is to state the obvious that LEPs should consider the lessons from the RDA evaluation experience and draw upon the evidence base that is currently available and will be available from initiatives such as the 'What Works Centre on Local Economic Growth'.

1.6 Research aims

This thesis tests the applicability of, and the factors which determine the success of, Evidence Based *Regional* Policy Making in England. An interdisciplinary approach across the health and regional policy sectors is taken. There are two major research aims:

Research aim 1: A dominant perspective within the EBPM literature has been to use EBM approaches as a yardstick against which wider social policy is assessed and to draw parallels between the practices of EBM and EBPM. There has not been such a study exploring the extrapolation of an EBM approach to the regional policy context.

This research aims to investigate how various types of evidence/knowledge are used across contexts and with different actors to understand what can be deduced about the generation, communication and use of regional policy evaluation evidence.

Research aim 2: The utilisation of knowledge translation tools and decision support tools as a mechanism to support the uptake of evidence is an emerging topic in the EBPM literature.

This research aims to critically analyse the role of a decision support tool to extend an EBM approach to regional policy investment prioritisation decision making.

These two research aims are expressed in terms of three research questions:

1. What are the epistemological and applicability challenges of extending an Evidence Based Medicine approach to regional policy evaluation?

This will be explored through a comparative analysis of the methodological guidelines and the central 'pull' for evaluation evidence for investment prioritisation across the health and regional policy sectors. The purpose is to reveal debates around evidence types and the role of research credibility.

2. What factors influenced the generation, communication and use of evaluation evidence within the English RDAs?

This will be explored through analysis of the perspectives of an expert stakeholder group to understand the application of the regional policy evaluation guidelines and the central 'pull' for evaluation evidence within the RDAs. The purpose is to reveal debates around how evidence was incorporated into the policy making processes of the English RDAs and what the role was of other factors besides evidence.

3. What are the potential opportunities and barriers to using a knowledge tool to extend an EBM approach to regional policy investment prioritisation?

This will be explored through analysis of the perspectives of an expert stakeholder group to understand the potential opportunities and barriers to the use of a decision support tool and to extend an EBM approach to regional policy investment prioritisation. The purpose of this question is much more normative than the first two questions and includes exploration of how to increase the uptake of evidence.

1.6.1 Adapting to the research context

The research began at the start of the Coalition's administration when details of its local growth policy were vague and there was a paucity of documentary evidence to rationalise Coalition thinking. The discipline of evaluation had also, in effect, been dismissed with evaluation budgets being one of the first areas of spending to be cut during austerity measures and RDA abolition. Therefore, the decision to undertake a comparative analysis of evaluation and investment prioritisation processes across health and regional policy was based upon the *conjecture* that evaluation would once again become a relevant government concern for regional/local growth policy, with the need to effectively prioritise investment and demonstrate the effectiveness of public spending. In some ways austerity measures placed an even greater emphasis on the need to employ effective policy evaluation.

The changing research context required the focus of the research to be adapted in two ways. Firstly, due to the context of RDA abolition and staff redundancies over a short time frame, the study population became hard to reach. Highly innovative research methods were developed to overcome these barriers, generating original research

management insights on the use of such methods. Secondly, as the policy context is changing so rapidly with constant new publications, I have provided a clear framework within which the latest ideas under the Coalition for the Local Enterprise Partnerships (LEPs) can be placed in comparison to the previous Labour RDA approach. Therefore, although regional policy has subsets and related fields (such as small business policy, skills policy and infrastructure planning), and a number of subnational institutions deliver economic policies; the scope of this research focuses on regional policy as delivered through the RDAs.

1.6.2 My place vis-à-vis the research

My motivation to carry out this piece of research primarily came from my own reflections transitioning from a research economist role within the health sector to evaluating economic development interventions within an RDA. As a health economist, I was tasked with working on health economics and outcomes research. This involved working on cost effectiveness models and technology appraisals to determine the quality of life impact of new therapies in comparison to cost for submission to NICE²² and other agencies. The work of NICE is highly controversial and has come under fierce criticism. Indeed, I experienced some of the practical implications and frustrations of these reported issues, particularly due to lack of transparency in the research commissioning process. However, overall I was encouraged by the culture of using evidence to inform commissioning and clinical decisions and the focus on improving evaluation methodologies to meet complex research questions.

When I began my role working as an evaluation analyst within an RDA I was struck by the methodological issues involved in analysing and attributing impact within the complex, highly political regional policy setting and by the relative paucity of guidance. As I started work, evidence was being collated for the 2009 national impact report on RDA spending (PWC 2009a). As the report was finalised, methodological concerns were highlighted both within and across evaluations, with the challenge of synthesising and aggregating the evidence base to evaluate RDA spending overall illuminated. In response to this, the updating of evaluation guidance (IEF + as it became known) (BIS 2009a) opened up a range of discussions within evaluation practice on the suitability of

²² The National Institute for Health and Care Excellence

economic evaluation methodologies and the role of evidence within the decision making process for regional policy. The change of government and axing of RDA budgets led to the need to be able to prioritise investment urgently within each agency and across the policy area. Although political ideology was at the forefront of these decisions, in my opinion, there was not an accessible and comparable repository of shared evaluation evidence to feedback effectively into this policy process.

While being a research economist within health and regional policy was important in enabling me to see the research questions raised by this study, in the following account the intention is to generate, not validate, understanding on the issues surrounding the use of evidence within public policy making.

1.6.3 Thesis overview

The following paragraphs give an overview of the structure of the thesis, leading the way into the succeeding chapters.

Chapter 2 introduces the theoretical background for this study. A review of the three conceptual questions in the EBPM literature are discussed: what kinds of evidence are used and the role of research credibility ('what'); how is evidence incorporated into policy making ('how'); and what are the other factors besides evidence which affect the way policy is made ('other factors'). In addition the theoretical roles of knowledge tools and decision support are reviewed. Chapter 3 describes the methodology, and details why the research was planned and carried out in this way. Chapters 4-6 then provide a comparative analysis of the methodological guidelines and central 'pull' for evidence drawing upon the academic and policy literature to address the first research question: what are the epistemological and applicability challenges of extending an Evidence Based Medicine approach to regional policy evaluation? The overarching EBPM debates are explored focusing on a comparative study of health policy and EBM (Chapter 4) and regional policy and impact evaluation (Chapter 5). Mirroring this investigative process within both sectors highlighted key differences across the sectors. The epistemological and applicability implications of extending an EBM approach to regional policy evaluation are analysed in Chapter 6.

A mixed-method sequential approach was then taken to explore the views of expert stakeholder groups involved in RDA policy evaluation. This included an online survey,

the development of a decision tool and an online workshop to test the applicability of such a decision tool to the regional policy context. The following chapters are driven by these empirical findings. Chapter 7 gives background, contextualising information on the policy making processes of the RDAs. Chapter 8 then presents the findings from the survey to address the second research question: what factors influenced the generation, communication and use of evaluation evidence within the English RDAs? Chapter 9 presents the online workshop findings, addressing the final research question: what are the potential opportunities and barriers to using a knowledge tool to extend an EBM approach to regional policy investment prioritisation? Chapter 10 presents the discussion, conclusions and recommendations of the study, reflecting upon the key findings in relation to the existing literature.

Chapter 2

Review of the EBPM literature

2.1 Introduction

To analyse Evidence Based Policy Making (EBPM) debates across policy domains, a conceptual framework was developed to structure examination of the existing literature and theories. In particular, a paper by Sutcliffe and Court (2005) was drawn upon which distinguished between three conceptual questions in the EBPM literature: what kinds of evidence are used and the role of research credibility ('what'); how is evidence incorporated into policy making ('how'); and what are the other factors besides evidence which affect the way policy is made ('other factors'). Similar themes were also surfaced by Rutter, Hawkins and Parkhurst (2013) when reviewing the Knowledge Transfer and Exchange (KTE) literature including debates centring on: the generation (production) of evidence; the communication (dissemination) and use (uptake) of evidence; and the contextual factors which influence the process. The three cross-cutting conceptual questions (what, how, other factors), in addition to the potential role of a knowledge tool to support EBPM, will be examined theoretically within this chapter. There is some degree of overlap between these cross-cutting debates. For instance, attempts to explain the types of evidence that are generated and are deemed credible necessarily engages with how such evidence is incorporated into policy processes.

2.2 Literature review methods

There is a wide body of literature focusing on the role of evidence in the policy making process. To explore the key themes of the generation and use of evaluation evidence, the topics of Evidence Based Policy Making (EBPM), Evidence Based Decision Making (EBDM) and the 'What Works' agenda in public policy were reviewed. Initially electronic and database searches of studies published in academic journals were conducted during October 2010-October 2011, including Google Scholar, EconLit and the World Wide Web of Political Science Abstracts. Other principal secondary sources included government department documents and professional/think-tank reports. The

reference list of each article or policy report was then reviewed to find additional articles. Given the vast quantity of EBPM literature, a snowballing method was then used for the literature review, utilising references from key papers in the field.

The literature review was an iterative process. For instance, when investigating the role of 'other factors besides evidence' it was found that there was a need to go beyond the EBPM literature to review parts of the political science literature to understand the political nature of evidence and decision making. In addition, during the literature review it was found that there is some ambiguity in the EBPM literature regarding the term 'evidence.' Contandriopoulos et al. (2010) point out that the terms 'evaluation', 'evidence' and 'research' have been used interchangeably with the term 'knowledge.' Moreover, it was found that there has been a marked shift in the EBPM literature towards the examination of 'knowledge' (Jones 2009) in recognition of a more holistic approach.

It emerged there was a need to explore the interrelation between the supply and demand of evidence to understand factors influencing evidence utilisation. Therefore the literature on 'Knowledge Transfer and Exchange' (KTE), defined as the interactive interchange of knowledge between research users and researcher producers (Kiefer et al. 2005) was also reviewed. This terminology is used by Mitton et al. (2007) and Rutter, Hawkins and Parkhurst (2013). Others have defined this as 'research utilisation' (Weiss 1979) and 'knowledge utilisation' (Ottoson 2009). It was found that acronyms and terminology used in this field vary, but terms used in this thesis include 'knowledge translation tools', 'knowledge management' and 'knowledge brokering.' These terms will be explained in due course and a paper by Estabrooks et al. (2006) was drawn upon to clarify terminology and theories in the field. So, although this thesis primarily focuses on the technocratic concepts of 'evaluation' and 'evidence', the broader concept of 'knowledge' is also an integral component.

2.3 Debate 1: Evidence types and the role of research credibility

The following review of the literature discussed below centres on evidence types and the role of research credibility ('what evidence'), alongside the similar themes of the 'generation' of evidence identified in the KTE literature. This predominantly provides the theoretical background for Chapters 4-6.

In terms of the generation and credibility of evidence, the EBPM/KTE literature critically reflects on the production of evidence and focuses on the conceptual questions of what should count as evidence for policy making, who should govern (or steer) the use of research evidence for policy and what is 'good evidence' for decision making (Rutter, Hawkins and Parkhurst 2013). Staunch advocates of EBPM argue the need for more scientifically rigorous research (including social science research) that has been systematically gathered, critically appraised and rationally analysed (Davies 2004; Sutcliffe and Court 2005). However, EBPM literature also suggests a more nuanced approach is needed given that policy makers, analysts and wider stakeholders may have different viewpoints on what types of evidence are most relevant and credible (Glasby and Beresford 2006). Nutley, Walter and Davies (2007) add that this is compounded by individuals being trained in different disciplines and thus holding different traditions.

Pawson et al. (2003) categorised five 'knowledge types': organisational; practitioner; user; research; and policy community knowledge. Sanderson (2003, p339-340) proposes the differing roles of 'episteme' (theoretical academic and research knowledge/evidence) and 'techne' (instrumental professional and institutional experience), as well as 'phronesis' (intrinsic virtues embodied in human practices during decision making). Similarly, Jones (2009) categorises three types of knowledge: participatory (civil society); research; and project and program. Project and program knowledge are described as encompassing experiential knowledge as well as evaluation processes. This distinguishes policy evaluation from external research and academic output.

Overall, the abundance of potential evidence available and the variety of methods by which evidence may be presented (e.g. "expert knowledge, published research, existing statistics, stakeholder consultations, previous policy evaluations, the Internet, outcomes from consultations, costings of policy options, output from economic and statistical modelling" (Cabinet Office 1999a, p33)) alongside the multitude of potential study designs (e.g. randomised clinical trials, systematic reviews, qualitative case studies, theory based evaluations etc.) compounds *uncertainty* over what constitutes rigorous, reliable and relevant evidence (Nutley 2003; Nutley, Walter and Davies 2007).

In terms of the supply of evidence, criteria have been developed in the social sciences literature to make judgements about the *rigour* of the evidence base. For quantitative research this traditionally includes the assessment of: internal validity; external validity; reliability; and objectivity. For qualitative research, criteria based upon the work by Guba (1981) and Lincoln and Guba (1985) includes the assessment of: credibility; transferability; dependability; and confirmability (these are analogous to the quantitative criteria). This is summarised in Table 1 below.

Table 1: Criteria for judging quantitative and qualitative research

Aspect	Traditional Criteria for Judging Quantitative Research	Alternative Criteria for Judging Qualitative Research
Truth value	Internal validity: the extent to which variations in an outcome or dependent variable can be attributed to controlled variation in an independent variable.	Credibility: the credibility criteria involves establishing that the results of qualitative research are credible or believable from the perspective of the participant in the research.
Applicability	External validity: inference that the presumed causal relationship can be generalised across alternate measures of cause and effect and across different types of persons, settings and times.	Transferability: the degree to which the results of qualitative research can be generalized or transferred to other contexts or settings. Transferability is enhanced by describing the research context and the assumptions that were central to the research.
Consistency	Reliability: consistency of a given inquiry is generally a precondition for validity. It refers to a study's consistency, predictability, dependability, stability and/or accuracy. Reliability typically rests on replication.	Dependability: emphasizes the need to account for the ever-changing context within which research occurs and how these changes affected the study.
Neutrality	Objectivity: neutrality, a demonstration that the inquiry is free of bias, values and/or prejudice.	Confirmability: refers to the degree to which the results could be confirmed or corroborated by others.

Source: adapted from Guba (1981, p80)

As such, study design has come to the fore as the key marker of the strength of evidence. Evans (2003, p78) notes “it has long been recognised that not all research designs are equal in terms of the risk of error and bias in their results.” Explicit ‘hierarchies of evidence’ have been developed, primarily within the EBM literature,

placing randomised experiments with clearly defined controls (RCTs) at or near the top followed by other less 'rigorous' approaches, with case study reports usually at the bottom (Petticrew and Roberts 2006; Bagshaw and Bellomo 2008). Moving beyond the EBM paradigm, a similar approach has been to develop categorisations of 'hard' objective versus 'soft' subjective types of evidence (Marston and Watts 2003); and Sefton (2000, p26) has developed a kind of hierarchy for economic evaluation²³ with quantitative evidence such as cost-benefit analysis at the top and qualitative evidence at the bottom.

Thus in EBPM literature, experimental research and quantitative data have been used as a benchmark to compare against non-experimental research and qualitative data. An important debate has focused not only on the development of such 'evidence hierarchies' but on their *applicability*, both within and across policy domains. Jones (2009) counsels the need to incorporate a wide breadth of evidence, or more holistically 'knowledge', into policy making. Despite this, Sutcliffe and Court (2005, p3) contend that policy makers make "hierarchical judgements" in choosing what evidence to use (and therefore demand), and argue that these judgements are often "deeply embedded in assumptions over validity and power." As such, they contend that the demand for evidence, and tendency to focus on a "limited range of 'top-end' evidence such as empirical research, policy evaluation and expert knowledge ...thereby creates an implicit evidence hierarchy."

Moving on to the demand for evidence, explicit criteria have therefore been developed to make judgements about the policy *relevance* of the evidence base and the feasibility of translating evidence into policy. For instance, bringing together the supply and demand perspectives for the design of economic policy evaluation, Sefton et al. (2002) identified key assessment criteria, including quantitative and qualitative considerations. The authors' narrative is summarised in Table 2.

²³ The most common definition of economic evaluation, and the one used in this thesis, is a "systematic attempt to identify, measure and compare the costs and outcomes of alternative interventions" (Sefton et al. 2002, p7 citing Drummond et al. 1997 and HM Treasury 1997). It is acknowledged that economic evaluation is part of the wider discipline of evaluation and synonymously that evaluation is only one type of evidence that could be used in the policy making process.

Table 2: Assessment criteria for the design of economic policy evaluation

	Concept	Quantitative	Qualitative
Validity	Validity refers to whether a study is able to scientifically answer the questions it is intended to answer.	Validity may depend on the extent to which the method used provides an unbiased estimate of outcomes and cost.	Validity may depend on how faithfully a study conveys people's experiences of an intervention.
Generalisability	This is whether findings can be generalised beyond the specific study population and setting.	Quantitative evaluators focus on how statistically representative their study sample is of the target group as a whole.	A broader definition would include other types of evidence, including knowledge about why certain interventions work in certain circumstances and for certain groups of people. This links to work by Pawson and Tilley (1997).
Relevance	This is about how useful the results are to decision makers (i.e. the extent to which the evidence can be translated into policy and whether the policy implications of the research are feasible and affordable). This links to work by Nutley, Davies and Walter (2002, p4) who identify the 'timeliness' of evidence as important	Most economic evaluation studies are designed to answer a specific question: whether the intervention being evaluated is an efficient use of resources, compared with alternative ways of using these resources.	Decision makers may be interested in a much wider set of evaluation questions, such as how to improve a programme.
Feasibility	This is the extent to which the proposed evaluation strategy can be implemented in practice, given the constraints on evaluators.		

Source: adapted from Sefton et al. (2002, p35)

When considering the policy relevance of evidence it is apposite to note Nutley's (2003) work on 'bridging the policy/research divide', Patton's (2012) work on 'utilization-focused evaluation' and Markusen's (2015) paper on problem-driven research in Regional Science. Taking a more holistic approach and embracing the broader concept of 'knowledge', some scholars have focused on 'knowledge-for-

action' theories in evaluation (Ottoson 2009). This links to the practicalities of translating evidence into policy. For instance, Shaxson (2005) advocated that a proportional approach should be taken to an evidence based approach given inevitable time and resource constraints. Colby et al. (2008) identified the need for: clear translation, accessible and easy-to-use information, and relevance to the policy context. Rutter, Hawkins and Parkhurst (2013, p9) note that one of the most common themes in such literature is the "call for knowledge outputs to be relevant to policy maker needs: fitting outputs to policy makers' timescales and agendas, and ensuring that the information output is relevant to the problem being solved."

2.4 Debate 2: The way in which evidence is incorporated into the policy making process

The following review of the literature centres on the way in which evidence is incorporated into the policy making process ('how'), alongside the similar themes of the 'communication' and 'use' of evidence identified in the KTE literature. This provides theoretical background for Chapter 8.

In terms of the communication of evidence, the KTE literature tends to focus on the conceptual questions of how research knowledge is typically translated into policy, how to improve the use or uptake of evidence in policy making and contextual factors which influence the process, often described as 'barriers' or 'facilitators' (Rutter, Hawkins and Parkhurst 2013).

One of the most common themes in the KTE literature echoes the section above on evidence types; the call for evidence to be policy/decision-relevant. "Packaging, translating, spreading and commissioning research are... strategies which have been developed in response to the overwhelming quantity of research evidence and its lack of relevance to decision makers" (Ward, House and Hamer 2009, p270). The accessibility of findings is a prevalent theme and studies of communication through print and electronic media and personal, face-to-face contact have highlighted that passive dissemination is ineffective (Kerner 2006; Grimshaw et al. 2006). In particular, knowledge management models have been developed in response to the difficulties associated with "navigating, managing and sharing a large body of research and other evidence" (Ward, House and Hamer 2009, p269). Another theme in the KTE literature

is the role of capacity building to address shortcomings in the ability of decision makers to access, interpret and apply research evidence (Ward, House and Hamer 2009). For instance, Court and Young (2003) note the importance of capacity building to ensure receptivity to research findings. Ward, House and Hamer (2009, p272) argue that “a more positive way of viewing the capacity building model is in fostering self-reliance in both the researcher and the decision maker, developing the knowledge transfer and communication skills and developing the analytical and interpretive skills of decision makers.”

A final consideration for the communication of evidence is the relationship between evidence producers, and users, incorporating the role of ‘knowledge brokers’ and their associated activities of developing positive relationships. Mitton et al. (2007) and Court and Young (2003) argue the importance of involving evidence users early in the research process to increase engagement and the uptake of evidence. Ward, House and Hamer (2009, p270) discuss the role of ‘knowledge brokers’ to act as intermediaries or linkage agents between evidence users and producers “to stimulate knowledge exchange, the development of new research and the application of solutions.” They note that knowledge brokerage can reside in individuals, organisations or structures (Ward, House and Hamer 2009, p268). Shaxson et al. (2012) highlight potential roles for individuals to: compile information; disseminate/translate ideas; link/network/facilitate; and to collaborate/manage relationships and processes. Similarly, Mitton et al. (2007) reviewed five frameworks developed to guide the process of KTE and highlighted the barriers and facilitators to the communication of evidence. In their review, the most important determinants of research utilisation were the mechanisms linking researchers and research users.

Focusing on the use of evidence in policy making processes, questions have tended to focus in the EBPM/KTE literature around what is the ‘good use’ of evidence from a governance perspective (Rutter, Hawkins and Parkhurst 2013). Supporters of EBPM advocate a more rigorous and systematic approach to policy making incorporating the use of evidence to inform and guide decisions. Yet the outcome of incorporating evidence into the policy making process is not widely understood (Rutter, Hawkins and Parkhurst 2013). There have been attempts to understand the utilisation of evaluation more widely within central (NAO 2010; Rutter 2012; BIS 2014b) and local government

(Percy-Smith et al. 2002; Allen, Grace and Martin 2014), but not within regional policy. Although not widely cited, Huber's work (2006) is apposite here. He constructed a set of functions, or roles that evaluation evidence can take within regional policy organisations (focused on EU regional policy instruments), including evaluation being used as: 'window dressing', a 'formal exercise' or part of a 'co-ordinated learning process.'

Research and other sources of evidence may be used in indirect and subtle ways within policy processes, however. Ottoson (2009) argues that a commitment to using knowledge in its original form cannot be achieved in real world policy processes and Mitton et al. (2007) suggest that success measures ought to focus on how the information was used rather than whether it was used. Ottoson (2009) considers that the aim of KTE can either be considered as 'top down change' where the knowledge dictates the policy or 'bottom up change' where the knowledge shapes the policy within the wider process. On a similar line, Armstrong and Wells (2006b, p265) note that there is sharp disagreement over the application of evaluation findings in policy and practice, with those who feel that evaluators should impose evidence informed change if needed at one end of the spectrum and those who feel that evaluators do not have this mandate and instead "should facilitate actors to reach a deeper understanding of what they are doing" at the other.

The EBPM literature contains a diverse range of studies which attempt to theorise and describe the use of evidence within the policy process. In a seminal paper, Weiss (1979) presented six models (see Table 3) to describe the various ways in which research can influence policy making: the knowledge driven; problem solving; interactive; political; tactical; and enlightenment models.

Table 3: How will policy makers use evidence?

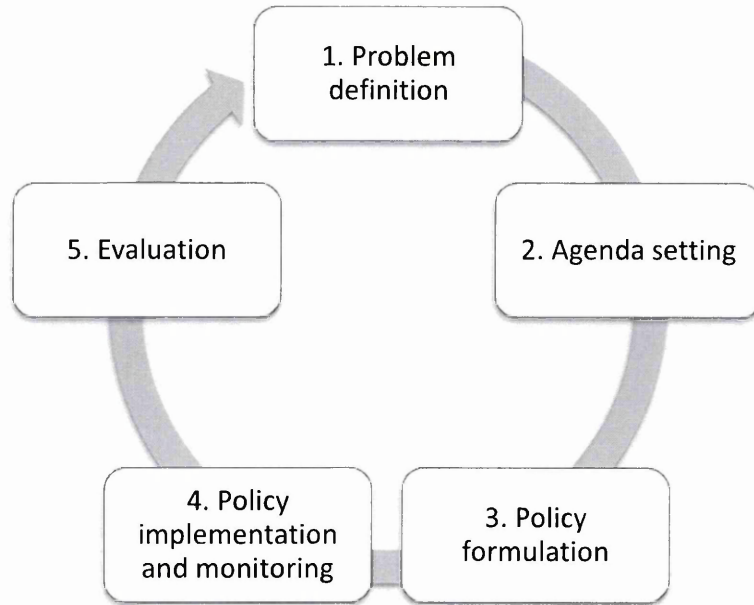
The knowledge driven model	This derives from the natural sciences. The fact that discoveries have been made sets up pressures for the development and use of the knowledge.
The problem solving model	This involves the direct application of the results of a specific study to an impending decision.
The interactive model	Here, researchers are seen as one set of participants among many. The use of research forms part of a complicated process that might also depend upon experience, political insights and pressures, social technologies and guesswork.
The political model	Here, research is used as political ammunition, especially where it is deployed to support a predetermined position.
The tactical model	Research may be used as a delaying tactic in order to avoid taking responsibility for unpopular policies or potentially negative outcomes.
The enlightenment model	This stresses the indirect influence of research rather than the direct impact of particular findings in the policy process. Thus the concepts and theoretical perspectives that social science engenders pervade the policy making process.

Source: Nutley and Webb (2000, p30 citing Weiss 1979)

These models propose that research may be used in a range of deliberate ways, for strategic or political ends, or to find a solution to a technical problem. Building upon this work, Jones (2009) identifies knowledge/policy ‘paradigms’ termed ‘rational’, ‘pluralism/opportunism’ and ‘politics/legitimation’. Literature focused on each of these three paradigms will be discussed in turn in the remainder of this and the following section.

Within the rational paradigm “knowledge is seen as providing instrumentally useful and essentially ‘neutral’ inputs that serve to improve policy, and policy making works in ‘problem-solving’ mode, according to logic and reason” (Jones 2009, p5). Reviewing the stream of literature based towards the ‘rational choice’ and ‘rational up to a point’ end of the spectrum, a key line of enquiry has been to clarify policy process frameworks and to identify the potential role of evidence for each discrete stage of the policy making process. In particular, a common approach is to use ‘policy cycles’ to present the discrete phases of policy making (as shown in Figure 2). Sutton (1999, p5) notes that such studies may align with a “linear model of policy making, characterised by objective analysis of options and separation of policy from implementation.”

Figure 2: An example policy cycle



Source: Adapted from Sutcliffe and Court (2005, p5)

Sutcliffe and Court (2005, p6) simplified the functions of the policy process into four categories and outlined some specific issues regarding the use of evidence for each stage. This work has been adapted in Table 4. The implication drawn from Table 4 is that different types of evidence are often needed for different parts of the policy process. Despite this, literature aligned with the 'rational' paradigm has inevitably come under criticism for taking a simplistic and naive empiricist view of the role of evidence in public policy.

Table 4: Components of policy processes and different evidence issues

Stage of the policy process	Description	Different evidence issues
Problem definition and agenda setting	Awareness and priority given to an issue.	The evidence needs here are in terms of scoping the issue, identifying new problems (or opportunities) or the build-up of evidence regarding the magnitude of a problem so that relevant policy actors are aware that the problem is indeed important. The political nature of evidence is to be discussed further in section 2.5.
Policy formulation	Determining the policy options and then selecting the preferred option.	Options analysis including understanding the instrumental links between an activity, output(s) and outcome(s) as well as the expected cost and impact of an intervention. It may be necessary to carry out research to provide new evidence.
Policy implementation and monitoring	How policy is put into practice and monitoring an intervention.	Here the focus is on operational evidence to improve the effectiveness of initiatives. A key factor is interpreting and applying evidence. Sound monitoring mechanisms need to be developed.
Evaluation	Assessing the process and impact of an intervention.	Evaluation is built upon sound monitoring mechanisms. Evaluation examines the actual outcomes and impact of policy, whether it meets its implicit objectives and what unintended impacts it has had and upon whom. Evaluation should determine the effectiveness of the implemented policy. A key factor is that evaluation should be communicated to provide the basis for future decision making through the continuing policy process.

Source: adapted from Sutcliffe and Court (2005, p6)

When focusing on the role of evaluation evidence specifically, the processes of appraisal, monitoring and evaluation, and the importance of evidence on policy effectiveness, are commonly emphasised in the literature (HM Treasury 2011, p14). However, it could be argued that such evidence has the potential to influence all stages of the policy cycle. For instance, the 'Enlightenment Model' (Table 3) suggests that evaluation may be able to influence the problem definition and agenda setting stage by challenging new ideas, providing new perspectives and reordering the policy-agenda. It could be argued that for strategic decision making at the problem

identification stage, this is about identifying need (rationale, resource impact, policy importance) but also about the factors which could impact on the feasibility of delivery and the likelihood of policy having an impact. At the policy formulation stage, it could be argued that the systematic use of evaluation evidence could be used to inform the allocation of investment between policy/programme areas, providing evidence of what works and why.

In addition, the linear policy cycle depicted in Figure 2 and Table 4 has been revised to account for the point that evaluations do not have to occur after a policy has been implemented (HM Treasury 2011, p15). Ex-ante evaluations, for instance pilots and trials, can be used to directly inform the policy development process. However, Pawson (2002, p1) argues that usually in practice “the policy cycle revolves quicker than the research cycle, with the result that 'real time' evaluations often have little influence on policy making.” This raises the issue of whether or not evidence is “timely” to be practically relevant to policy processes (Nutley, Davies and Walter 2002, p4).

The pluralism/opportunism paradigm “challenges assumptions about the rationality of the policy process, seeing it as involving pragmatic decisions taken based on multiple factors in the face of uncertainty. The incorporation of knowledge involves often erratic and opportunistic processes, and explicit efforts of various actors” (Jones 2009, p5). Reviewing the stream of literature based towards the ‘decision making in disorder’ end of the spectrum, a key line of enquiry has been to clarify that policy makers are not *able* to make decisions rationally. Pressman and Wildavsky (1984) argue that there is often a ‘deficit’ between policy formulation and implementation, particularly as the number of elements in the policy process increases. Clay and Schaffer (1984) argue that policy formulation and implementation are best understood as a “chaos of purposes and accidents” (cited by Sutton 1999, p32) and Cohen, March and Olsen (1972) refer to decision making as an anarchic ‘garbage can’ process with solutions looking for problems rather than the other way. Of particular relevance is Schön's work (1983) that highlighted the complexity of the policy environment and the political nature of evidence. This leads on to the next debate.

2.5 Debate 3: Other factors besides evidence which affect the way policy is made

The following review of the literature centres on other factors besides evidence which affect the way policy is made ('other factors'), alongside the theme of 'factors which influence the KTE of evidence' identified in the KTE literature. This also provides the theoretical background for Chapter 8.

Following on from the discussion in the section above, and moving on to the politics/legitimation paradigm of Jones' work (2009, p5), Rutter, Hawkins and Parkhurst (2013, p14) note "this focuses on the role of power in the policy process, analysing the impact of actors, norms, institutions and discourse." This embraces that policy makers do not necessarily *seek* to make decisions rationally. When considering the role of other factors besides evidence, there is a tendency for the EBPM/KTE to focus on the way evidence has been promoted to justify policy decisions, rather than inform or guide them, emphasising the political nature of decision making and the (instrumental) selection of evidence. This links to the previous discussion of Huber's (2006) work on evaluation evidence being used as 'window dressing,' giving a veneer of credibility to policy processes externally.

Lavis et al. (2003) note other factors besides evidence affect the way policy is made including: the orientation of the governing party/supporters; stakeholder views; public opinion; who wins/loses from policy; decision making rules; and past policy (institutions/path dependency). However, the political science literature emphasises the ways that evidence itself cannot address political issues and values (Rutter, Hawkins and Parkhurst 2013). Greenhalgh and Russell (2006, p35) argue that policy making is not actually concerned with what works and is rather about "making and implementing collective ethical judgments" which aim to pursue "the right course of action in a particular context, at a particular time, for a particular group of people and with a particular allocation of resources." Likewise Abeyasinghe and Parkhurst (2013) highlight that evidence alone cannot say what the policy 'should' be, given a backdrop of multiple social concerns at stake. The term 'Evidence Based Policy Making' has therefore been criticised for suggesting that there are technocratic solutions to what are, essentially, political problems (Rutter, Hawkins and Parkhurst 2013).

It has also been argued that the nature of knowledge itself can affect its uptake. Contandriopoulos et al. (2010) investigated the link between the 'technocratic' understanding of issues and consensus, and discuss the concept of 'issue polarisation.' They argue that in cases of 'low level issue polarisation' technically focused decision making can be employed with rational dialogues and arguments, based upon a similar world-view amongst actors. This is linked to Schön's (1983) work on 'high, hard ground problems' which are either more amenable to technical understanding or are less important to both individuals and wider society. Contandriopoulos et al. (2010) argue that in cases of 'high level issue polarisation' there are more likely to be political debates and a strategic approach towards knowledge use. Likewise, Schön (1983) used the metaphor of a 'swamp' to describe important, complex, and messy problems that resist technical analysis. In turn, Rutter, Hawkins and Parkhurst (2013) make the link that the nature of knowledge can therefore lead to bias in policy making processes towards policies for which there is already nearly universal agreement. Drawing upon Schön's work, Parsons (2002, p45) notes that policy makers must "chart a course [through the swamp], navigating the hindrances to their progress and engaging not just with facts but values and politics." Rutter, Hawkins and Parkhurst (2013, p 17) argue that with this understanding, "there is a need to shift from a scientific-rationalist frame" and that "rather than a narrow focus on what works, the alternative is to consider what is appropriate in the circumstances, and given the overall policy objectives."

Sanderson has done much work on the manifestations of EBPM which fail to acknowledge the fundamentally political nature of evidence. He contends that although EBPM should focus on policy learning (i.e. understanding which types of intervention work best and why), there is an inherent bias in EBPM towards certain types of knowledge (i.e. that which is "derived through quantitative methodologies, empirically-tested and validated") which is seen to lead to "instrumental rationality" and a managerialist and mechanistic approach to policy making (2002, p6). This links to the previous discussion of Huber's (2006) work on evaluation evidence being limited to the minimum required for a purely 'formal exercise' rather than as part of a 'coordinated learning process.' Similarly, Rutter, Hawkins and Parkhurst (2013) discuss the concept of 'Evidence Controlled, Managed and Legitimised Policy' (ECMLP) rather

than evidence based policy. Therefore, in recognition of the crucial role of ideas and ideology in EBPM and the discursive nature around evidence in policy debates, Sanderson (2002) encourages greater reflexivity and deliberation by policy makers. Drawing all these lines of argument together, Wells (2007, p27) concludes that:

EBPM is only one component of the policy making process. Ideas, values, political strategies and previous practice are probably of greater significance. However... evaluative research undertaken with an understanding of political ideas, institutions and contexts provides a richer basis on which to inform policy, and equally, practice.

2.6 Debate 4: The role of knowledge translation tools and decision support

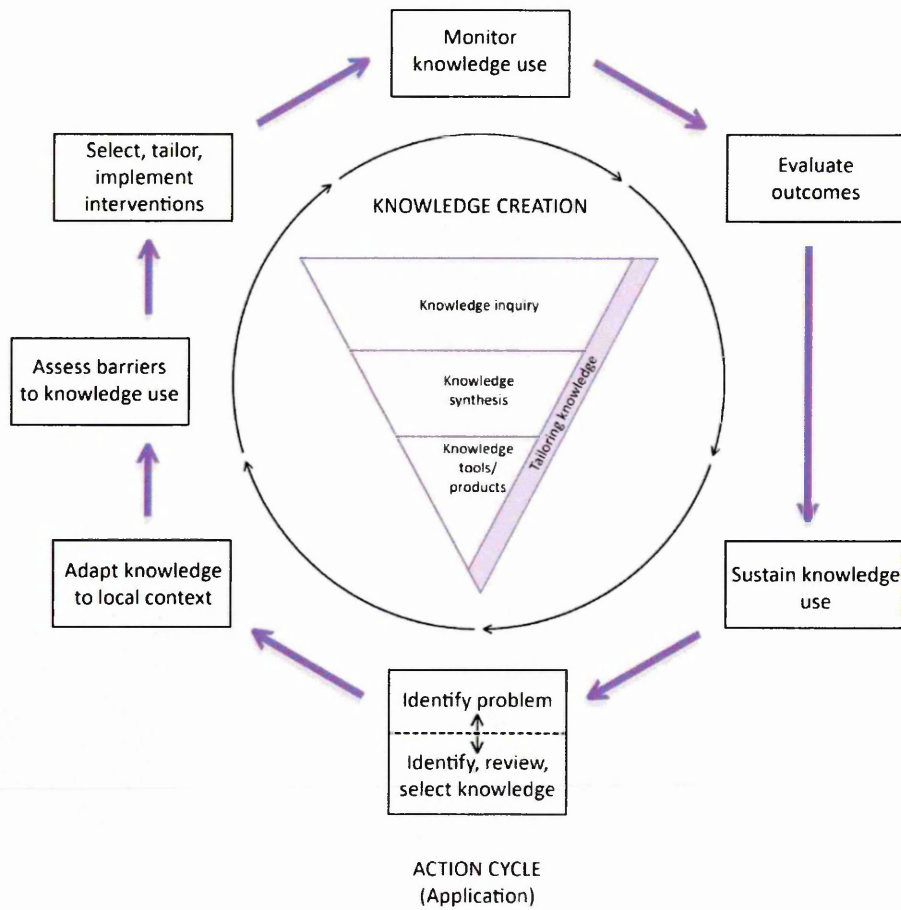
The following review of the literature centres on the role of knowledge translation tools and decision support. This predominantly provides the theoretical background for Chapter 9.

The role of knowledge translation tools and decision support essentially aligns with a strand of the EBPM/KTE literature focused on how to achieve the maximum impact for a body of evidence and the strategies that may be employed. Graham et al. (2006, p19) developed a model of the knowledge translation process (Figure 3) detailing the theoretical steps of creating, tailoring and applying knowledge. In this framework a 'knowledge creation funnel' and a 'knowledge action cycle' are illustrated. It is conceptualised that as knowledge moves through the funnel, it becomes more distilled and refined and presumably more tailored to the needs of stakeholders. Then this knowledge feeds into the start of the action cycle at the 'identify problem' stage (Straus, Tetroe and Graham 2013). The action cycle then describes the implementation or application of knowledge. Although an abstraction of a complex and dynamic process, of interest is that the authors' stress that research knowledge must be translated into forms conducive to policy maker engagement to increase research utilisation.

Knowledge from primary studies is referred to as 'first generation knowledge' and knowledge synthesis is referred to as 'second generation knowledge' (Straus, Tetroe and Graham 2013). They note that approaches for drawing knowledge from the

aggregation of existing knowledge include: systematic reviews; meta-syntheses; scoping reviews; and realist reviews (Straus, Tetroe and Graham 2013). Borenstein et al. (2009) also discuss the role of meta-analysis, which is a statistical approach to provide a pooled estimate measure of effectiveness for comparison between interventions/studies building upon a systematic review of the evidence (Glass 1976). In the EBPM/KTE literature, such reviews are proposed as a means of bridging the gap between research and decision making, going beyond the raw data (Lavis et al. 2003). 'Knowledge translation tools' are termed 'third-generation knowledge', and in Figure 3 the role of knowledge tools/products is an integral component of this 'Knowledge to Action Framework' (Graham et al. 2006, p19).

Figure 3: Knowledge to Action Framework



Source: Graham et al. 2006, p19

Knowledge translation tools could take the form of a decision support system (DSS), which draw upon a repository of information (knowledge-base) and utilise an inference mechanism (logic). The role of decision support is an emerging debate within the EBPM/KTE literature. This is curious given that, in practice, decision making is often messy and complex (Cohen, March and Olsen 1972). Within individual or collaborative decision making scenarios, there will be elements of 'known' and 'unknown' information, alongside changes in the decision making context and elements of risk and uncertainty (March 1982). March claims that decision makers are constrained by cognitive capabilities (1994) and cognitive biases (1978). Recognising and understanding these factors are challenges that face decision makers.

An exploratory literature review was conducted focusing on the central debates surrounding cognitive processing for decision making and the factors that may bias the process. McCaughey and Bruning's (2010) work within health care was found to be apposite as they make the link between cognitive errors in human decision making and the implications this has for the assumptions of rationality underpinning EBPM approaches. Indeed, by its nature, evidence based policy decision making assumes a degree of individual rationality and utility maximization on the part of individual decision makers (Lin 2003).

McCaughey and Bruning (2010) go on to dispute the inherent assumption in EBPM that policy makers are capable of accurately analysing decision information, are resistant to influences and biases, and seek to make decisions that maximise societal benefit. By studying cognitive information processing and decision making they make the point that individuals are unique in terms of their personalities, abilities, beliefs and values. They argue that a decision maker's utility is highly subjective, open to the influence of affect (i.e. the experience of feeling or emotion) and may include variables such as personal gain, risk tolerance, relevance to related events, and value of a decision to the organisation. They argue that, analytically, individuals will interpret and assimilate data in different ways and at different speeds, even when the same data is apparently available to all. They highlight the role of 'heuristic' errors, defined as potential intuitive processing errors, whereby simplifications, or mental 'rules of thumb' and 'gut feeling responses' may undermine an evidence based approach. The point is also made that decision makers each have different life experiences and political beliefs, thus people will rank individual and social gains differently. Finally, they consider that group decision making has to take account of such individual objectives and biases, in addition to the processes of persuasion and opinion influencing (see also Bazerman 1998).

This once again links to Sanderson's work (2003, p339-340) and the differing roles of 'episteme' (theoretical academic and research knowledge/evidence) and 'techne' (instrumental professional and institutional experience), as well as 'phronesis' (intrinsic virtues embodied in human practices during decision making). Overall, this work highlights the complexity surrounding who is involved in decision making processes, how they participate and the context of the decision making environment.

Due to the large number of considerations involved in many decisions, approaches to decision support have emerged to overcome the constraints of analytically processing information to inform and guide decision making, termed *decision support systems* (DSS) by Gorry and Scott Morton (1971). DSS may be computer based and offer support to decision making processes, for instance, through simulating aspects of a decision scenario alongside summarising the evidence base. Generally, DSS approaches draw upon expertise from a wide range of disciplines including quantitative analysis, information systems and cognitive psychology (French, Maule, and Papamichail 2009). The DSS literature is extensive but an overview is provided by Pervan and Arnott (2005) and a survey of DSS applications is provided by Eom and Kim (2006) demonstrating applications across a range of sectors, including health care, operations and finance. Simulation modelling can also refer to the use of Cost Benefit Analysis models and Sunstein (2000) notes that CBA can assist decision making directly through comparing cost-benefit ratios between alternative actions but also indirectly as a form of challenge to long held beliefs, values and views.

Utilising a decision support tool as a strategy to support an EBPM approach leads back to considerations of the purpose of the communication and use of evaluation evidence discussed earlier (section 2.4). Nonetheless, Rutter, Hawkins and Parkhurst (2013, p18) argue that “most works discussing evidence uptake fail to engage with the political nature of decision making, or to critically assess the relevance of a given body of evidence” (section 2.5).

2.7 Conclusion

This chapter has introduced the theoretical background for this study. The approach taken and literature review methodology has been described. Three conceptual questions in the EBPM literature have been discussed: what kinds of evidence are used and the role of research credibility (‘what’); how is evidence incorporated into policy making (‘how’); and what are the other factors besides evidence which affect the way policy is made (‘other factors’). In addition, the theoretical roles of knowledge tools and decision support have been reviewed. The following chapter will describe the methodology for the research.

Chapter 3

Methodology

3.1 Introduction

This chapter describes the methodology, and details why the research was planned and carried out the way it was. The research context, methodologies chosen to investigate the research questions, and the sampling framework will be discussed. The methods for each phase of the study are then presented in detail, including discussion of empirical data collection, analysis and interpretation. Overall lessons from the research management process are surfaced and reflected upon in the conclusion.

3.2 Rationale for the methodology

The methodological framework and research strategy were designed with the intention of choosing the best methods to meet the aims of the research. The questions to be explored and the methods employed are:

1. What are the epistemological and applicability challenges of extending an Evidence Based Medicine approach to regional policy evaluation?

This will be explored through a comparative analysis of the methodological guidelines and the central 'pull' for evaluation evidence for investment prioritisation across the health and regional policy sectors. The purpose is to reveal debates around evidence types and the role of research credibility.

2. What factors influenced the generation, communication and use of evaluation evidence within the English RDAs?

This will be explored through analysis of the perspectives of an expert stakeholder group to understand the application of the regional policy evaluation guidelines and the central 'pull' for evaluation evidence within the RDAs. The purpose is to reveal debates around how evidence was incorporated into the policy making processes of the English RDAs and what the role was of other factors besides evidence.

3. *What are the potential opportunities and barriers to using a knowledge tool to extend an EBM approach to regional policy investment prioritisation?*

This will be explored through analysis of the perspectives of an expert stakeholder group to understand the potential opportunities and barriers to the use of a decision support tool and to extend an EBM approach to regional policy investment prioritisation. The purpose of this question is much more normative than the first two questions and includes exploration of how to increase the uptake of evidence.

3.2.1 Practical considerations

The empirical research was undertaken at a particularly challenging political, and economic, point in time. As discussed in Chapter 1, the formation of the Coalition government in May 2010 led to a dramatic and rapidly evolving change of policy direction, to austerity cuts and to the eventual abolition of the RDAs by March 2012. This context created a situation whereby the study population became hard to reach and required new thinking about how best to recruit participants and to approach the research as a whole. In addition, the removal of a regional tier of government also had the potential to undermine the credibility of the research when approaching and involving participants. The discipline of evaluation had, in effect, been dismissed, with evaluation budgets being one of the first areas of spending to be cut during austerity measures and RDA abolition. Therefore, the decision to undertake a comparative analysis of evaluation and investment prioritisation processes across health and regional policy was based upon the *conjecture* that evaluation would once again become a relevant government concern for regional/local growth policy, with the need to effectively prioritise investment and demonstrate the effectiveness of public spending. In some ways austerity measures placed an even greater emphasis on the need to employ effective policy evaluation. Therefore, a decision was made to explicitly acknowledge RDA abolition and recent events within the research.

3.2.2 Philosophical considerations

In terms of the theoretical perspective and underlying assumptions of the research, “recognition of the importance of an in-depth understanding of context and the diverse viewpoints of stakeholders” (Bryman 2012 citing Greene 1994, 2000) was identified within the literature review. A plurality of perspectives has therefore been

embraced to explore the shades of opinion across individuals, using appropriate methodologies, and given the practical constraints resulting from RDA abolition. A major tenet of applying such a pragmatic approach is that quantitative and qualitative methods are compatible and that research practices lie somewhere on a continuum between the two (Newman and Benz 1998). The comparative literature review, survey and workshop aimed to build understanding of the factors that are important for the generation, communication and use of evaluation drawing upon the EBPM/KTE theory, whilst embracing that the nature of evidence and decision making are politically charged and value laden, drawing upon political science theory.

3.3 Methodological framework

A mixed methods approach was taken (Creswell and Clark 2011) incorporating quantitative (numerical data) and qualitative (text data) methodologies, with the aim of each illuminating the other (Dixon-Woods et al. 2005). Neither method on its own was sufficient to capture both the underlying trends in the data as well as the rich detail of participants' perspectives. As presented in Table 5, the core aspects of the research strategy were a comparative literature review, online survey, the development of a knowledge tool and an online workshop.

Table 5: Methodological framework

Phase		Procedure	Product
Scoping phase		<ul style="list-style-type: none"> Scoping telephone interviews Desk based research 	<ul style="list-style-type: none"> Refined research questions Triangulation check for internal consistency
Phase 1: Comparative literature review		<ul style="list-style-type: none"> Desk based research 	<ul style="list-style-type: none"> Data display tables Key findings
Phase 2: Online survey	Data collection	<ul style="list-style-type: none"> Cross-sectional web-based survey Purposeful/snowball sampling 	<ul style="list-style-type: none"> Quantitative data Qualitative data
	Data analysis	<ul style="list-style-type: none"> Data screening Frequencies Coding and thematic analysis 	<ul style="list-style-type: none"> Descriptive statistics Data display tables
	Data interpretation	<ul style="list-style-type: none"> Explanation of the meaning of quantitative analysis Interpretation of the meaning of qualitative analysis 	<ul style="list-style-type: none"> Key findings
Phase 3: Knowledge tool		<ul style="list-style-type: none"> Development of a knowledge tool based upon dummy RDA data 	<ul style="list-style-type: none"> Knowledge tool to be used for elicitation in the online workshop
Phase 4: Online workshop	Data collection	<ul style="list-style-type: none"> Online workshop Purposeful/snowball sampling 	<ul style="list-style-type: none"> Quantitative data Qualitative data
	Data analysis	<ul style="list-style-type: none"> Data screening Frequencies Coding and thematic analysis 	<ul style="list-style-type: none"> Descriptive statistics Data display tables
	Data interpretation	<ul style="list-style-type: none"> Explanation of the meaning of quantitative analysis Interpretation of the meaning of qualitative analysis 	<ul style="list-style-type: none"> Key findings

The findings from each phase of the study were not intended to replicate each other, but to provide a different perspective on the issue. Both the survey and the online

workshop methods relied on the responses of individuals from the same overall study population and themes relating to all research questions were surfaced by all phases of the research. The findings from the survey directly shaped the design of the online workshop, however, and triangulation methods were employed to compare results across and between research phases.

3.3.1 Sampling framework

The aim was to recruit an expert stakeholder group with background knowledge and insights which were of direct relevance to the research questions. The goal was to elicit responses from commissioners and producers, as well as from users of evaluation evidence, across the policy cycle. The study population included RDA officers (including ex-RDA officers), central government officers from the Department for Business, Innovation and Skills (BIS) and the Department for Communities and Local Government (CLG) and external evaluators who had worked on RDA evaluations.

Initially, the research strategy chosen for Phase 2 (survey) was to employ a quantitative approach to statistically analyse the relationship between variables (Porter and Carter 2000, p19). The aim was to obtain a probability sample that would produce valid findings and which had a claim to be representative of the wider study population (Bryman 2012, p187) (i.e. the whole population of individuals working across RDA policy evaluation processes). However, a large sample would have been needed to permit statistically significant discriminatory variables to be determined, and to draw statistical inference with the required precision (Ritchie, Lewis and Elam 2003).

As highlighted by Table 6 below (column 2), the target study population became “hard to reach” (Bryman 2012, p418) due to RDA staff redundancies. In June 2010 the Government announced that the RDAs were to be abolished by 31 March 2012. However, the specific difficulty for conducting this research (which commenced in October 2010) was the speed of the transition to closure, the volume of redundancies early on in the process through ‘voluntary’ schemes and the different timings of the redundancy schemes across the RDA network. The contact details of many potential participants quickly became inaccessible and RDA Human Resource (HR) departments were unable to provide details of onward employment placements or to provide

accurate staffing figures for RDA departmental sub-groups. It became apparent it was not going to be possible to obtain a probability sample or to adjust or weight data given the absence of a sampling frame.

Therefore, a purposive, sequential, sampling approach was taken to sample participants in a “strategic way” (Bryman 2012, p418).²⁴ Individuals were intentionally selected who were “information rich” (Patton 1990, p169) and who worked (or had worked) in evaluation, strategy, economic appraisal and performance management roles. Attempts were made to obtain contact details for potential participants by: reviewing the websites of each of the RDAs/CLG/BIS; reviewing the business networking site LinkedIn; reviewing the websites of external consultancy firms; and contacting HR departments for each of the RDAs/CLG/BIS. However, this strategy proved fruitless. Potential participants were then targeted by contacting gatekeepers (managers and evaluation personnel) at each of the nine RDAs as well as at BIS and CLG. The contact details of these evaluation personnel were known due to my involvement with the cross-RDA evaluation network.

Ideally for Phase 2 the gatekeepers would have provided all the contact details (of current RDA staff, onward contact details for personnel who had already left the RDA, and external consultancy personnel) so that a response rate could be calculated. However, sometimes the approach taken was for the gatekeepers to directly email out the survey to their relevant contacts. The reason for this was twofold. Firstly, as many RDA and external consultancy personnel had already been made redundant, their contact details were not disclosed for reasons of confidentiality. Secondly, it was felt that it was an opportunity for gatekeepers to stress the importance of the research to their known contacts to increase the response rate. For Phase 4, gatekeepers (managers and evaluation personnel) were once again contacted. “Sequential snowball sampling” was also adopted to recruit additional participants when an opportunity presented itself and participants proposed others who had the experience or characteristics relevant to the research (Bryman 2012, p424). Table 6 presents further detail on the (approximate) target study population and actual sample sizes for the main organisations represented in the survey and online workshop.

²⁴ A discussion of the strengths and weaknesses of this sampling approach are discussed in section 3.6 for the online survey and section 3.8 for the online workshop.

Table 6: Sampling framework

Participant's organisation	Approx. number of RDA staff (FTE)* ²⁵	Approach to evaluation**	Survey participants (number and percentage)	Workshop participants (number and percentage)	Participants in both survey & workshop
RDA: • Yorkshire Forward (YF)	2010-11: 363 2011-12: 153	YF mainly took a project-level approach	48 (59%) ²⁶	-	-
RDA: • East Midlands Development Agency (emda)	2010-11: 228 2011-12: 48	emda took a programme level approach, using a sampling framework	1 (1%)	3 (16%)	1
RDA: • One North East (ONE)	2010-11: 324 2011-12: 121	ONE mainly took a project-level approach	6 (8%)	8 (42%)	2
Central government: • Department of Business, Innovation and Skills (BIS) • Department for Communities and Local Government (CLG).	-	BIS was responsible for co-ordinating evaluation across the RDA network and commissioned a number of national and sub-national evaluations	3 (4%)	8 (42%)	2
Other RDAs	-	-	10 (12%)	-	-
External consultancy	-	-	13 (16%)	-	-
TOTAL (N)			81	19	

Source: * YF 2011a, p31; YF 2012, p12; emda 2012, p73; ONE 2012, p46. ** Survey findings (see Chapter 8).

For the workshop, given that the knowledge tool to be used in the workshop was initially developed with (limited) consultation with YF personnel, a decision was made

²⁵ Full Time Employee (FTE).

²⁶ Majority RDA sub-group.

to target participants from former emda officers, ONE officers and central government officers. These organisations (cases) were chosen due to their differing approaches to evaluation, to give a more nuanced picture (Table 6). ONE had taken a similar approach to evaluation as YF and it was therefore possible to triangulate the findings across the online survey and workshop to check for internal consistency. Given that 5 of the workshop participants had also completed the online survey, 95 unique users, producers and commissioners of evaluation evidence contributed towards the research.

3.4 Initial scoping interviews

Initial exploratory and informal scoping telephone interviews were conducted in February 2011. Such enquiry helped to shape ideas on the topic, the problems and potential of theories and methods, and signposted to secondary data sources. This helped to refine the research questions and to shape the early stages of designing the empirical research.

3.5 Phase 1: Comparative literature review

A comparative study of the generation and use of evaluation evidence for investment prioritisation was conducted across the health and regional policy sectors drawing upon the methodological guidelines for evaluation and the academic and policy literature. As Collier (1993, p105) notes “comparison is a fundamental tool of analysis. It sharpens our power of description, and plays a central role in concept-formation by bringing into focus suggestive similarities and contrasts among cases.” The aim of the review was to investigate how various types of evidence/knowledge are used across contexts, sectors and with different actors focusing on the generation, communication and use of regional policy evaluation evidence and drawing out the epistemological and applicability implications of extending an EBM approach to regional policy evaluation. The approach taken is an interpretive ‘contrast of contexts’ (Skocpol and Somers 1980) to examine the two sectors (cases), to highlight their differences and thus interpret how parallel EBM processes for evaluation and investment prioritisation are played out in different ways within each context.

When reviewing the health care/ medical sector, it was found that the topics of health policy, Evidence Based Medicine (EBM) and evidence-informed health policy are well

researched and integrate several related disciplines including: epidemiology; biostatistics; behavioural sciences; health economics; healthcare management; and health knowledge transfer and exchange. The potential body of literature for consideration was vast and therefore it was decided to explore some key themes that emerged from the literature that was of particular relevance to the first research question. Search criteria were formulated by reviewing health economics text-books (Culyer and Newhouse 2000; Drummond et al. 2005) and methodological guidelines (NICE 2013b). This provided a starting point for mapping the field. Electronic and database searches of studies published in academic journals were conducted in an iterative manner during October 2010-October 2014, including searches of Google Scholar, MEDLINE and CINAHL. Government policy documents were also retrieved from sources such as NICE and the Department of Health. A key paper that was recently published from the perspective of NICE discussing the Institute's approach to the development of social policy was also reviewed (Alliance for Useful Evidence 2014). The reference list of each article or policy report was reviewed to find additional articles.

When reviewing the regional policy sector, it was found that the topics of Regional Policy, Evidence Based Regional Policy Making (EBRPM) and regional policy evaluation integrate several related disciplines including: regional studies/science; social and economic development/regeneration; spatial economic analysis; policy analysis; and policy evaluation. EBRPM is an under-researched area in comparison with the vast body of literature written on EBM. Therefore the objective was to explore some key themes that emerged from the multidisciplinary literature. Search criteria were formulated by mirroring the analysis undertaken for health policy. Electronic and database searches of studies published in academic journals were conducted in an iterative manner during October 2010-October 2014, including Google Scholar, EconLit and the World Wide Web of Political Science Abstracts. Specific journals were targeted such as 'Regional Studies'. Other principal secondary sources included government department documents from BIS,²⁷ CLG,²⁸ DG REGIO,²⁹ NAO,³⁰ and ONS³¹ alongside

²⁷ The Department for Business, Innovation and Skills (BIS).

²⁸ The Department for Communities and Local Government (CLG).

²⁹ Directorate-General for Regional and Urban Policy of the European Commission (DG REGIO).

RDA reports, and professional/think-tank reports. The reference list of each article or policy report was reviewed to find additional articles.

To undertake the comparative analysis, studies were used which have drawn parallels between the practices of EBM and EBPM (Dobrow, Goel and Upshur 2004; Cookson 2005) and which used EBM approaches as a yardstick to measure against wider, social policy evaluation (Sefton 2000; Sefton et al. 2002; Sefton 2003; Somekh et al. 2005). Given that such an approach has not been applied within the regional policy context, the aim was to elaborate and refine these theories by undertaking a structured analysis across the health and regional policy contexts.

3.6 Phase 2: Online survey

A self-completion survey (questionnaire) was chosen to collect information on participants' views and pre-existing knowledge (Bryman 2012, p231-243). Surveys incorporating a strong quantitative element tend to involve systematic questioning using mostly closed questions (Bryman, p249), whereby all participants are asked consistent questions to yield data which is standardised (Sapsford 1999). The aim of standardisation is to produce research findings which are representative of the population being researched, however, the reliability of such findings is dependent upon the sampling approach (May 1997).

Given the challenges of RDA abolition affecting the sampling frame, a decision was made to incorporate a strong qualitative element to the survey. Usually other methods, such as qualitative interviews, lend themselves better to exploring, in-depth, the perspectives of individuals and the context within which they operate (Bryman 2012, pp468-498). However, a web-based survey methodology was chosen as the most feasible way of canvassing a range of views from individuals involved in RDA regional policy evaluation quickly during the time of RDA abolition and as a strategy to collate quantitative data to capture underlying trends. The ability of the qualitative approach to allow participants to discuss the subject 'on their own terms' was seen as an important factor. Therefore, free text comments (open questions) were used to enable

³⁰ National Audit Office (NAO).

³¹ Office for National Statistics (ONS).

non-standardised responses and allow scope for new ideas to be introduced beyond the pre-set questions (Bryman 2012, p246).

Prior to the research being conducted, to the researcher's knowledge, survey methodology had not previously been used to investigate the views of an expert stakeholder group involved in RDA regional policy evaluation. The National Audit Office (NAO) had formally reviewed the evaluation functions of the RDAs previously, via an Independent Supplementary Review (ISR) process, however, this was based upon "a review of documents, observation of routine meetings, site visits, and tailored interviews/ focus groups with internal and external stakeholders" rather than via survey methodology (NAO 2010, p41). Since the survey was administered, however, the NAO conducted a web survey of chief analysts and analysts (n=15) between July 2012 and March 2013 to gather quantitative data on how evaluation evidence is used in practice and how it has contributed to policy decisions across 17 departments in government (2013, p46). This formed part of the evidence base for the report 'Evaluation in Government' which was quoted in Chapter 1 (section 1.4). However, this survey did not focus solely on departments which are responsible for spatial policy and did not encompass a strong qualitative element.

3.6.1 Data collection

As no previous surveys on this topic were found at the time of conducting this research, data was collected through a bespoke survey (see Appendix 2 for the full survey). The survey was hosted on the SurveyMonkey website (www.SurveyMonkey.com). As a (previously) practising RDA evaluation officer I had some prior understanding of what elements of content and structure might be used in the survey. A literature review was also conducted to inform the development of the survey items incorporating a brief exploration of organisation theory literature, focused on strategic decision making (SDM). The importance of context for SDM was underlined by Papadakis and Barwise (1997, p291) and, in terms of shaping the survey items directly, the literature review surfaced particular contextual factors to consider. In the paper by Papadakis, Lioukas and Chambers (1997), the importance of managerial, organisational, external environmental and group dynamics factors (citing Schneider and DeMeyer 1991) as well as decision-specific factors (citing Pettigrew 1990) were examined. In particular, this shaped survey item 17(a) (Appendix 2): 'what

strategic, operational and relational challenges do you think the evaluation team faced?', and shaped the predetermined series of statements to which respondents were asked to express agreement, disagreement or neutrality.

The survey was designed using a simple format including: self-assessment items measured on a 5-point Likert type scale; dichotomous answers like "Yes" and "No"; and open-ended questions. Some questions in the survey had an open-ended "Other (specify)" option. A choice of "Not applicable" (NA) was included when necessary. Questions requiring in-depth knowledge and experience were targeted at appropriate participants, determined by departmental sub-group. For instance, questions focused on the role of evaluation evidence for strategic decision making were targeted at RDA strategy teams, whereas questions focused on the role of monitoring processes were targeted towards RDA delivery and performance teams. The survey was designed to ensure that a maximum response rate was obtained for four core quantitative questions (N=81): the overall influence of evaluation; evaluation team processes; challenges faced by evaluation; 'value for money' of evaluation (Appendix 2). Further specific, quantitative and qualitative survey items were targeted at the following departmental sub-groups:

- RDA evaluation officers³²
- RDA delivery/performance officers³³
- RDA economic appraisal officers³⁴
- RDA strategy officers³⁵
- Central government officers
- External evaluation consultants

³² For instance in YF: current or previous personnel within the YF Strategy Directorate Evaluation Team.

³³ For instance in YF: current or previous personnel within the YF delivery directorates including: Business, Economic Inclusion, Environment and Finance.

³⁴ For instance in YF: current or previous personnel within the YF Strategy Directorate Chief Economist's Unit.

³⁵ For instance in YF: current or previous personnel within personnel within the YF Strategy Directorate Economic Policy and Strategy team (excluding evaluation team members).

The survey was pilot-tested with the PhD supervisory team and with the Evaluation Manager at Yorkshire Forward to secure content validity (Carmines and Zeller 1991, p20). Based on the pilot-test results, survey items and technical problems with the survey administration were revised accordingly.

Administering the survey

A range of approaches was used to maximise the response rate. Initially gatekeepers (managers) were contacted and asked to discuss the survey with their current (or past) team members. Potential participants were contacted via email or through the messaging service of the business-related social networking site, Linked In (www.linkedin.com). Contact was made via an introductory email, information was provided to build trust such as researcher background information, and confidentiality was emphasised. The web link (URL) to the survey directed respondents to the SurveyMonkey website. Clear instructions on how to complete the survey were provided at the beginning of the survey and the amount of personal information requested was limited. A three phase follow-up sequence was used (based upon Dillman 2007). To those subjects who had not responded by the set date (1) five days after distributing the survey URL, an e-mail reminder was sent out; (2) ten days later, the second e-mail reminder was sent; (3) two weeks later, the third e-mail reminder was sent stating the importance of the participant's input for the study. A communication tracker document was kept.

The survey was carried out in two phases. In phase 1 it was administered to RDA officers between May and June 2011 (during the time of the RDA abolition process). In phase 2 the survey was administered to external evaluators and central government officers from BIS and CLG between July and August 2011.³⁶

Ethical considerations

Ethical approval for the research was gained by completing a 'Proforma for Post-Graduate Student Projects' which was then scrutinised by the Head of Research Ethics

³⁶ 2 responses were collected between October 2011 and March 2012 from external evaluators.

at Sheffield Hallam University³⁷ to decide whether the research required full ethical review by the Faculty Research Ethics Committee (FREC). The proforma was submitted on the 26th July 2011 and was revised and resubmitted after feedback from the Head of Research Ethics. Ethical approval was gained on 12th September 2011.

An introductory email and a paragraph at the beginning of the survey detailed the participant information. A formal consent slip was not required, however, because completing the survey demonstrated consent and the involvement of participants was entirely voluntary. The confidentiality of information was maintained and the anonymity of participants respected. The research did not cause harm to the participants, involve sensitive topics, or involve vulnerable groups (Bryman 2012, p146-7).

Respondents

A total of 81 participants completed the online survey. As shown in Table 7, responses were received from 65 RDA officers (80%) (48 YF officers; 17 'other' RDA officers), 13 external evaluators (16%) and 3 central government officers (4%).³⁸ In Table 7, the RDA respondents are then presented by departmental sub-group. The dispersion of respondents across departmental sub-groups within the survey broadly reflects the dispersion within the RDAs (see Appendix 4), with most RDA staff employed within 'Delivery Directorates' and fewer staff employed within 'Strategy.' However, a greater proportion of Strategy Directorate personnel completed the survey. This was due to the sampling approach taken but probably an influential factor was that the subject was of interest and was more relevant to these participants (Edwards et al. 2002).

The sample was roughly balanced between evidence users and producers with 44% (36) of respondents³⁹ being placed within predominantly research focused roles. The sample was an experienced group with 58% (38) of the respondents from the RDAs having worked within their role for 5 years or more.

³⁷ Head of Research Ethics, Sheffield Hallam University. Email communication 14.05.2012.

³⁸ Appendix 3 presents a full breakdown of respondents by each RDA.

³⁹ Including RDA evaluation officers, RDA economic appraisal officers and external evaluation consultants.

Table 7: Demographics of respondents

	Number	Percentage of sample
Respondent's organisation		
RDA:	65	80%
• YF	(48)	(59%)
• Other RDA	(17)	(21%)
External consultancy	13	16%
Central government	3	4%
Respondent's department		
RDA:		
• RDA evaluation officers ⁴⁰	14	17%
• RDA delivery/performance officers ⁴¹	34	42%
• RDA economic appraisal officers ⁴²	9	11%
• RDA strategy officers ⁴³	8	10%
External consultancy	13	16%
Central government	3	4%
TOTAL (N)	81	

The RDA sample was biased towards Yorkshire Forward (YF) personnel in particular given that I already had access to participant contact details and was known to them as a bone fide doctoral researcher without the need for third party endorsement.

Reflections on the data collection

The timing of the research had a direct impact on validity. Potentially, either a quantitative survey approach could have been taken (with a large probability sample to generate statistically representative findings), or a qualitative interview approach could have been taken (with a sufficiently small sample to permit in-depth analysis and generate findings which are highly contextual and not intended to be generalised). However, given that the ability to contact the study population of interest was rapidly diminishing, a pragmatic and pluralist approach was taken to data collection. The survey enabled access to an expert stakeholder group with background knowledge and

⁴⁰ For instance in YF: current or previous personnel within the YF Strategy Directorate Evaluation Team.

⁴¹ For instance in YF: current or previous personnel within the YF delivery directorates including: Business, Economic Inclusion, Environment and Finance.

⁴² For instance in YF: current or previous personnel within the YF Strategy Directorate Chief Economist's Unit.

⁴³ For instance in YF: current or previous personnel within personnel within the YF Strategy Directorate Economic Policy and Strategy team (excluding evaluation team members).

insights which were of direct relevance to the research topic. The qualitative questions enabled exploration of views and attitudes of individuals (Given 2008) embedded across RDA policy processes. It was not the purpose of this research to produce findings which were representative of the entire study population for generalisation, but to yield data which was highly contextual and embedded within respondents' perceived reality (Given 2008). The qualitative data was richly productive of new ideas, and was used in a complementary way to 'explain' areas where the quantitative data had raised questions, and perhaps to 'fill in gaps' where the quantitative data had failed to give a sufficiently full picture. The survey methodology was also a successful strategy to collate quantitative results that indicate the underlying trends in the data and frame discussion of the qualitative data.

A further issue relating to validity was my known identity as a previous RDA evaluation officer, which may have influenced response to some questions and potentially may have influenced who responded to the survey. The opening paragraph of the survey contained an appeal to respondents to take part in the research, reinforcing my previous role. Although I endeavoured to make the language of the survey neutral, asking questions about the utilisation of research findings inevitably discloses that I was working from an ideology of evaluation evidence being useful to policy making. It should be remembered that "questions are live communications and different questions will convey different intentions of what it is that the researcher wants to the respondent" (Marsh 1979, p302). However, Marsh advises "we must not confuse an impossible attempt to achieve 'absolute truth' through asking unbiased questions, with the aim of being objective in our quest for truth" (Marsh 1979, p304).

Nonetheless, it is important to recognise that response bias is likely, given firstly the phrasing of some of the quantitative questions (in particular the predetermined series of statements) which may have led to issues of evocation and intensification; secondly, the respondent's relationship to the researcher and to evaluation team members; thirdly, the timing of the survey during RDA abolition meaning that stakeholders had highly vested interests; and finally, the potential desires of participants to provide 'textbook responses'. This may have led to overestimation of the potential role for evaluation and intensified responses on the role of other factors besides evidence in the policy making process. This will be reflected upon in Chapter 8.

3.6.2 Data analysis

The online facility (SurveyMonkey) collected the results accurately without researcher error. Data were then exported into Microsoft Excel and all statistical analysis of the quantitative results was conducted in that spreadsheet package. This was chosen over other packages such as SPSS (Statistical Package for the Social Sciences) as only limited statistical analysis was needed. Data screening included reviewing missing data and outliers.

Perhaps the main challenge for analysing the quantitative data was that departmental sub-group survey items were a key part of the survey design. However, sample sizes for some departments were very low, affecting validity (Table 7). Although all the data were analysed, a decision was made to focus the discussion towards data from the four core quantitative survey items to which all participants had responded (the overall influence of evaluation; evaluation team processes; challenges faced by evaluation; and 'value for money' of evaluation). Frequency analysis was conducted to identify percentages for responses to the questions in the survey (Appendix 5). 'Missing value' cases were reported in the discussion but were omitted from the analysis (Swift 1996).

Given that the RDA sample was biased towards Yorkshire Forward (YF) personnel in particular, analysis was undertaken to compare the responses from YF and 'other RDA' organisational subgroups. As analysis of the quantitative data did not suggest heterogeneity between sub-groups, the quantitative results are presented for the total population when reporting the survey findings in the narrative. For information purposes, the quantitative results are also presented for the YF sub-group through the use of footnotes so that the thread of the narrative is not compromised. For the qualitative data, quotations are presented denoting the departmental sub-group and identification number of the respondent. For reasons of anonymity (given the small sample sizes of departmental sub-groups), and given that analysis of the quantitative data did not suggest heterogeneity between sub-groups, the qualitative data across organisations has been merged and does not separate YF personnel from other RDA personnel.

Detailed responses were given to the qualitative questions with a total word count across the survey of over forty thousand words. A challenge of conducting a survey with a strongly qualitative element was that a lot of free text, therefore non-standardised, data was gathered. All free text comments were included in the analysis. Such in-depth analysis and detailed exploration of the qualitative data required intensive analytical work. Free text comments were highly diverse, ranging from comments on the subject of study to comments of the survey design, as well as views on the future of regional policy and evaluation. Theoretically, numbers of unfavourable and favourable responses could have been counted for common themes to infer overall agreement or disagreement on issues raised (Berelson 1952). In reality, this was not possible because there were so many shades of opinion which could not be clearly categorised and elements of positive/negative comments could co-exist in the same paragraph of free text. In addition, as cautioned by Krane, Andersen and Streat (1997, p214) “rare experiences are no less meaningful, useful, or important than common ones. In some cases, the rare experience may be the most enlightening one.”

Initially NVivo 9 qualitative data analysis software was used to categorise, order and analyse the data. However, it was found to be too rigid when refining matrices and themes. Therefore ‘framework analysis’ was used to facilitate the ordering and synthesising of the data as suggested by Ritchie, Spencer and O'Connor (2003, p219). For the qualitative thematic analysis, the steps included: preliminary exploration of the data units (statements, sentences, etc.); constructing an index of central themes and sub-themes which were then represented in a matrix (table); applying the matrix to the data by segmenting and ordering the text into the matrix; aggregating similar themes together; connecting and interrelating themes; and constructing a narrative. Overall, the process was highly iterative and involved refining the matrix and themes, and rereading the data multiple times. 11 themes and 49 sub-themes were then presented through visual data-display matrices (tables) presented Appendix 6 and drawn upon within the narrative using key quotations.

Reflections on the data analysis

Data analysis and interpretation is presented in Chapter 8. Ideally it would have been desirable to undertake sub-group analysis to compare the responses across

organisations and departmental groups. For quantitative data, procedures such as the chi square test can be used for comparisons between groups of respondents to give the probability of the relationship between variables, and statistical significance can be reported at different levels (for instance, using $p < 0.05$) (May 1997). However, such sub-group analysis could not be undertaken using such a procedure because of the small sample sizes involved and because the data were not selected through a probability sampling approach. Similarly, for the qualitative analysis it would have been interesting to conduct thematic analyses within and across contexts (organisations/departments) and cases (individuals) (Bryman 2012, p417). However, given the large volume of data generated by the survey approach, such analysis was deemed unfeasible given the scope of this research, but could form the basis of future work.

3.7 Phase 3: Decision support tool

The theoretical background for the rationale and potential use of a decision support tool is provided in Chapter 2 (predominantly in section 2.6). The aim of Phase 3 was not to produce a complete, exemplary decision support tool, but to produce a simple, prototype decision support tool which could be used for elicitation in the online workshop. Later in the thesis it is hypothesised that programming an EBM approach into the underlying model for a decision tool to support regional policy investment prioritisation, will reveal practical analytical and data access implications. In addition, drawing upon the EBPM/KTE literature review (Chapter 2), it could be hypothesised that use of such a knowledge translation tool has the potential to support an EBRPM approach and to increase the utilisation of evidence. These findings are reflected upon and interpreted in Chapter 9.

The following section gives an overview of the tool. Inevitably this leads to a certain level of discussion about the underlying model and data underpinning the tool. To fully understand the issues surrounding the construction of the decision tool requires a certain level of understanding about the policy making processes of the RDAs, their intelligence functions and Management Information Systems. In Chapter 7 the processes of an example RDA, Yorkshire Forward (YF), are described to give background, contextualising information to the construction of the decision tool and to position the findings of the online survey and online workshop. However, it is

necessary to include a section here giving a high level account of the design of the decision tool to frame the discussion of its use within the online workshop (Phase 4).

3.7.1 Developing the tool

I developed an initial prototype decision support tool whilst working at Yorkshire Forward (YF), the RDA for Yorkshire and the Humber. The decision support tool was developed in a Macro-Enabled Workbook in Microsoft Excel and it was designed to support RDA strategic investment prioritisation decision making and to allocate budgets across programmes and projects. The tool drew upon a knowledge base of evidence drawn from across the RDA directorates and employed a decision logic so that users would be able to simulate aspects of a decision scenario alongside summarising the evidence base (Chapter 2). The tool was simplified and dummy data developed for the purposes of the research. The simplified decision support tool and underpinning data were sense checked during the pilot testing of the workshop (see section 3.8).

Overall, the approaches taken to programme the simulation model and map, collate, clean, analyse and rank the data and the underlying assumptions governing such programming choices are integral, yet highly contentious, considerations in the construction of the tool. These key findings are examined further in Chapter 9. Nonetheless, when reflecting upon the decision tool presented in scenario 2 in the workshop, participants reflected that the decision tool appeared credible:

I would do this - and did work up a similar model [to scenario 2] at ONE. (ONE officer)

The scenario appears credible, as this was the sort of situation that was faced by the RDAs. (Central Government officer)

The simulation model

Essentially the tool worked by ranking projects in order of those that had higher benefits compared to costs, and it was programmed to enable a change in the total budget available (i.e. a change of the resource allocation threshold). Ideally the simulation model would have been based upon Cost Benefit Analysis (CBA) principles

(Mishan and Quah 2007) so that the expected benefits would be monetised and weighed up in comparison to the costs of each project to produce a cost benefit ratio.

For most projects it was possible to calculate a cost per output ratio. The cost per output ratio was used because, when planning the construction of the decision support tool, it was found that decision makers within YF placed emphasis on output data over GVA benchmark data (to be discussed in section 5.3.2). However, in practice it was found that RDA monitoring systems were not designed to produce cost per output data. A crude method employed for the decision tool was to interview senior managers to elicit the primary output (i.e. jobs created) for each programme area and to then survey project managers to elicit the proportion of investment that was to be spent by each project on activities to generate this primary output. This gave a crude cost per output which was then primarily used to rank projects against each other within the model. For projects focused on producing Strategic Added Value (SAV)⁴⁴ outcomes, a numeric scale was applied. A crude method employed was for project managers to identify the level of SAV produced using a traffic light system (high/medium/low SAV produced) so that these projects could be ranked against each other. Then the decision tool would equally split the total budget across programme types (i.e. skills, business assists, CO2 reduction, SAV etc.). If the budget were reduced, then it would be the best performing projects within each programme that the decision tool would suggest should go forward and the total costs and benefits of these projects would be presented.

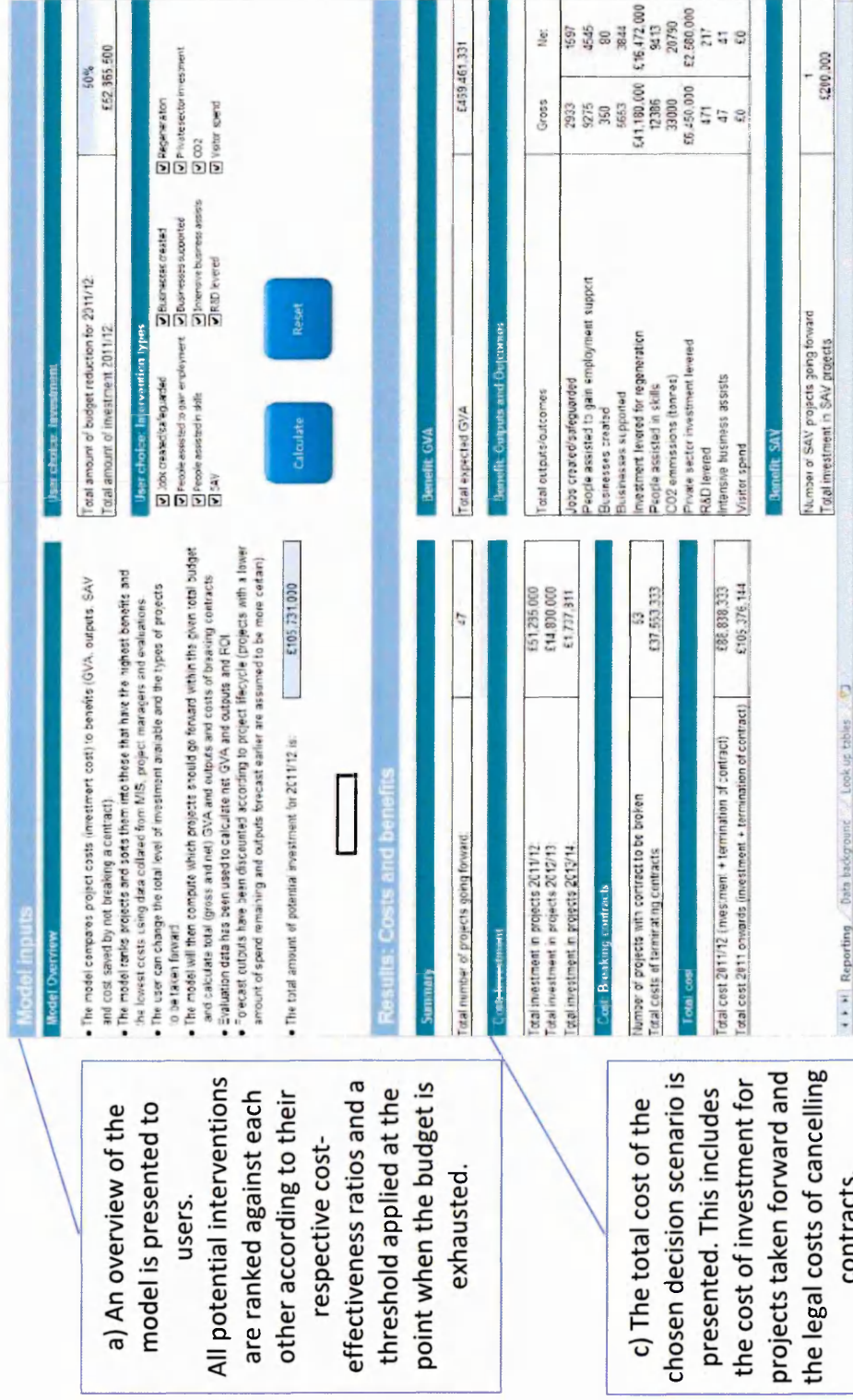
Clearly this programming was in the early stages of conceptualisation. Future programming would have needed to strengthen the allocation of the budget between intervention 'types,' mapping 'primary' outputs and outcomes more clearly on to potential future programme types. In addition, future programming could have enabled decision makers to place weightings (i.e. preferences) on the different programme areas (i.e. 20% of the budget to be allocated to CO2 reduction) rather than a simple binary response (yes/no) to taking forward certain intervention types.

⁴⁴ "Strategic added value was a concept that tried to encapsulate the role of RDAs in delivering unquantifiable benefits such as regional leadership and partnership working" (YF 2011b, p15) i.e. it was the role of influencing others to take action to meet regional and national objectives.

Reporting

A screenshot of the 'reporting' worksheet of the decision tool is presented in Figure 4. The key features are annotated and explained further in the discussion below. The model inputs (user choices) and results (total costs and benefits) of the decision tool are presented on one page to promote interactivity and so that the user is able to see the outcome of a chosen decision scenario easily and run alternative scenarios. Future programming could have enabled storage and retrieval functions to enable easy comparison.

Figure 4: Screenshot of the decision support tool presented in the workshop (user inputs and results)



a) An overview of the model is presented to users.
All potential interventions are ranked against each other according to their respective cost-effectiveness ratios and a threshold applied at the point when the budget is exhausted.

b) User input into the model is limited but users are able to change the resource allocation threshold to run budget reduction scenarios. Users are also able to only take forward certain intervention 'types' to align with policy direction and individual preferences.

c) The total cost of the chosen decision scenario is presented. This includes the cost of investment for projects taken forward and the legal costs of cancelling contracts.

d) The total estimated benefits of the chosen decision scenario are presented. These include total GVA, (gross and net) outputs and outcomes and an indication of the investment into projects focused on Strategic Added Value (SAV).

a) Model overview

In Figure 4, it is shown that an overview of the model is presented to users within the first part of the first worksheet of the decision tool, giving a very high level synopsis of the simulation model and underpinning data. However, the underlying assumptions and limitations of the tool and data are not presented here. An assumption was made that, given the level of analytical understanding and digital literacy needed to use the tool; further training and support would be needed to guide and inform individuals to use the tool (appropriately) in practice.

b) User choice

In terms of user input, the decision tool for the workshop was designed to be very simple with many of the calculations being performed 'behind the scenes.' Users are able to input to an extent. In Figure 4, it is shown that users are able to choose the resource allocation threshold (the total budget) and the tool was programmed so that users could easily run budget reduction scenarios (reducing the resource allocation threshold by 10%, 50% etc.) using a drop down box with pre-determined options. Users are also able to only take forward certain intervention 'types' (a proxy for programme types) to align with policy direction and individual preferences using checkboxes as a binary control function (yes/no). The decision tool then includes interventions which have a 'ticked' checkbox in the analysis and does not include interventions which have an 'unticked' checkbox. The decision tool enables users to input all of their choices before clicking on a 'calculate' button (command button) to execute the macro and run the decision scenario.

c) Results: costs

In Figure 4, below the 'user choices', the tool then presents the outcomes of the chosen decision scenario. Outcomes (results) are presented in terms of total benefits and costs. The total number of projects going forward is initially reported.

Then the cost of investment is reported, breaking down the profile of investment over the following three years (given that RDA projects were usually committed for 3 years). The level of investment was captured annually for capital and revenue spend and for Single Pot and other funding (both actual and forecast spend) within the Management

Information System. Next, the total number of projects to be cancelled is reported and calculation of the costs of breaking those contracts presented. RDA projects were usually committed for 3 years within legally binding contracts. The cost of terminating these contracts was based upon the proportion of investment that was expected to have to be repaid if a contract was broken. Finally, the total cost is presented including the cost of the investment in projects going forward and the costs of breaking contracts for cancelled projects.

d) Results: benefits

For presentation of the benefits, impact, outcome and output indicators are presented. Both gross and net data are presented (see Chapter 5 for the additionality calculation). Total expected GVA is presented first. Secondly, output and outcome indicators are presented aligning with the technical guidance on core outputs/indicators to be collected by the RDAs (See DTI 2006, p141). Outputs collated include: Jobs Created, Employment Support (assisting people to gain employment), Businesses Created, Businesses Assisted, Regeneration (investment levered for regeneration – Public/Private), Skills (people assisted in skills), Sustainable Development (carbon emission reduction), Private Sector Investment (private sector investment levered), R&D Leveraged, Intensive Assists (intensive businesses assisted). An additional outcome indicator, visitor spend, was added to enable the inclusion of tourism projects in the model analysis.

Finally, an indication of the investment into projects focused towards Strategic Added Value (SAV) is presented by reporting the number of projects going forward whereby SAV was the primary goal of the intervention. The total investment into SAV projects is presented (summed over the three year RDA spending profile).

Data background

A screenshot of the 'data background' worksheet of the decision tool is presented in Figure 5, the key features are annotated and explained further in the discussion below.

Figure 5: Screenshot of the decision support tool presented in the workshop (data background)

Project name	Project description	Local Authority	Actual Single Pot spend to date (capital + revenue)	Forecast Single Pot investment 2011/12	Forecast Single Pot investment 2012/13	Forecast investment
Forward together	Business support project to encourage business start-up and growth in Alton	Rotherham District	£700,000	£400,000	£200,000	
Media office space	Office space for creative and media industry	Sheffield District	£2,000,000	£500,000	£0	
Business crime	Advice on business crime	Whole region	£700,000	£250,000	£250,000	
R&D funding	R & D support scheme	Whole region	£2,500,000	£1,300,000	£3,000,000	£1
Training support	Funding for training for businesses	Whole region	£3,000,000	£4,000,000	£0	
Increase innovation	Innovation advisers	Whole region	£2,500,000	£1,400,000	£600,000	
Skills advice	Skills advice to businesses	Whole region	£3,500,000	£5,000,000	£5,000,000	
Digital funding	Early stage funding to digital and media business	Whole region	£200,000	£5,000,000	£0	
Conference centre	Build new conference centre	York	£50,000	£2,500,000	£0	
Business Recovery Advisory Service	Support for businesses to improve efficiency, raise productivity and develop new ideas	Whole region	£8,000,000	£500,000	£0	
Advanced manufacturing	Support programme encouraging inward investment in advanced manufacturing	Bradford District	£350,000	£500,000	£0	
SCSE improvement	Project to improve GCSE attainment of young people	Kirklees District	£4,000,000	£400,000	£0	
IT skills	Network for businesses in the IT sector supporting collaboration	Whole region	£200,000	£30,000	£0	
Venture capitalists network	Support towards the network of venture capitalists	Whole region	£1,000,000	£600,000	£0	
IT programme	Programme to encourage the use of ICT in businesses	Whole region	£1,000,000	£15,000	£0	
Women entrepreneurs network	Support for a network of female business owners and leaders	Whole region	£500,000	£120,000	£80,000	
Collaboration	Collaboration between universities and business	Whole region	£5,000,000	£100,000	£50,000	

a) Dummy data was produced based upon a prototype model I built at YF. Data was collated from: the RDA Monitoring and Information system; RDA project managers and RDA legal personnel; and from published evaluation data.

b) Dummy data was created for 100 hypothetical RDA projects.

In terms of the data underpinning the tool, a process of mapping internal and external data sources within YF was undertaken. This included brainstorming with key personnel in other teams supplying data and undertaking unstructured interviews with decision makers about relevant decision scenarios and model outputs. This initial tool was underpinned by data collated from 'Artemis' (the RDA's Management Information System(MIS)), from a survey with project managers, the Chief Economist's Unit, the contracting team, the legal team and from published benchmark evaluation data. Given issues of data access and confidentiality, for the purposes of the research dummy data was developed to feed into a simplified version of the tool.

a) Data sources

MIS data included: project name; programme type (YF 'policy product range'); directorate; project performance management status (YF 'Project Management Framework' stage); Local Authority; Funding type (Single Pot/ERDF split); Actual Single pot spend to date (capital and revenue); Forecast single pot spend profile for following 3 years (capital and revenue); Forecast gross outputs profile for following 3 years. The proportion of single pot remaining and timing of the forecast outputs were used to calculate a 'project cycle weighting.' In effect, this was to account for risk (i.e. projects near to completion were likely to be more certain to deliver forecast outputs compared to projects at the beginning of the project lifecycle).

Data collated from RDA senior management and project managers included: project description (open text box); project rationale (closed question with predetermined choices); intervention type, defined as the primary output type (closed question with predetermined choices); apportionment of project investment to the primary output (closed question with predetermined choices); expected SAV (closed question with predetermined choices); interdependency to other projects (closed question with predetermined choices); interdependency to other projects (open text box); other information (open text box). The predetermined choices for the closed questions were determined in consultation with other RDA evaluation team and strategy officers.

Contracting data collated from RDA senior management and project managers included: potential contract costs if the project is to be cancelled (closed question with predetermined choices); other information (open text box). The predetermined

choices for the closed questions were determined in consultation with RDA contracting and legal team personnel and included 'no costs', 'exit costs', 'previously ineligible costs' and 'settlement costs.'

Socio-economic data for each Local Authority collated from the Chief Economist's Unit included: Working Age Employment Rate 2009 (%); Claimant Count July 2010 (Working Age) %; Enterprise Starts per 10,000 Adults 2008; % Working Age Qualified to NVQ4; % Working Age Qualified to NVQ2; Equivalised Household Income 2009; and Average Household Income 2009. These data were indexed to the average for Yorkshire. Socio-economic data was not included in the simplified version of the model.

b) Hypothetical projects

Dummy data was created for 100 hypothetical RDA projects (interventions) which could have been plausibly funded by an RDA.

Look-up tables

A screenshot of the 'look-up tables' worksheet of the decision tool is presented in Figure 6, the key features are annotated and explained further in the discussion below.

Figure 6: Screenshot of the decision support tool presented in the workshop (look-up tables)

Look up tables									
<p>Business: Individual enterprise level support</p> <p>Business: Science R&D and innovation infrastructure</p> <p>Business: Inward investment</p> <p>Business: Sustainable consumption/production</p> <p>Business: Infrastructure development of indigenous business</p> <p>Business: Other</p> <p>Place: Bringing land back into use</p> <p>Place: Public realm</p> <p>Place: Mass transit and tourism</p> <p>Place: Other regeneration interventions</p> <p>People: Skills and workforce development</p> <p>People: Attracting people to jobs</p> <p>People: Supporting the development of educational infrastructure</p> <p>People: Provision of level 3 or above qualifications</p> <p>People: Hybrid assets</p> <p>Integrated Island</p>									
01 - NO OF JOBS CREATED OR SAFE GUARDED	02 - NO OF PEOPLE ASSIGNED TO EMPLOYMENT	03 - NO OF BUSINESSES CREATED	04 - NO OF BUSINESSES SUPPORTED	05 - NO OF PUBLICLY OWNED INVESTMENT LEVERAGED %	06 - NO OF SUSTAINABLE INVESTMENT LEVERAGED %	07 - NO OF PEOPLE ASSIGNED TO DEVELOPMENT	08 - NO OF PEOPLE ASSIGNED TO DEVELOPMENT	09 - NO OF PEOPLE ASSIGNED TO DEVELOPMENT	10 - PRIVATE SECTOR INVESTMENT LEVERAGED %
0.10	0.49	0.23	0.66	0.10	0.76	0.02	0.02	0.02	0.40
0.12	0.49	0.23	0.66	0.10	0.76	0.02	0.02	0.02	0.40
0.14	0.49	0.23	0.66	0.10	0.76	0.02	0.02	0.02	0.40
0.16	0.49	0.23	0.66	0.10	0.76	0.02	0.02	0.02	0.40
0.18	0.49	0.23	0.66	0.10	0.76	0.02	0.02	0.02	0.40
0.20	0.49	0.23	0.66	0.10	0.76	0.02	0.02	0.02	0.40
0.22	0.49	0.23	0.66	0.10	0.76	0.02	0.02	0.02	0.40
0.24	0.49	0.23	0.66	0.10	0.76	0.02	0.02	0.02	0.40
0.26	0.49	0.23	0.66	0.10	0.76	0.02	0.02	0.02	0.40
0.28	0.49	0.23	0.66	0.10	0.76	0.02	0.02	0.02	0.40
0.30	0.49	0.23	0.66	0.10	0.76	0.02	0.02	0.02	0.40
0.32	0.49	0.23	0.66	0.10	0.76	0.02	0.02	0.02	0.40
0.34	0.49	0.23	0.66	0.10	0.76	0.02	0.02	0.02	0.40
0.36	0.49	0.23	0.66	0.10	0.76	0.02	0.02	0.02	0.40
0.38	0.49	0.23	0.66	0.10	0.76	0.02	0.02	0.02	0.40
0.40	0.49	0.23	0.66	0.10	0.76	0.02	0.02	0.02	0.40
0.42	0.49	0.23	0.66	0.10	0.76	0.02	0.02	0.02	0.40
0.44	0.49	0.23	0.66	0.10	0.76	0.02	0.02	0.02	0.40
0.46	0.49	0.23	0.66	0.10	0.76	0.02	0.02	0.02	0.40
0.48	0.49	0.23	0.66	0.10	0.76	0.02	0.02	0.02	0.40
0.50	0.49	0.23	0.66	0.10	0.76	0.02	0.02	0.02	0.40
0.52	0.49	0.23	0.66	0.10	0.76	0.02	0.02	0.02	0.40
0.54	0.49	0.23	0.66	0.10	0.76	0.02	0.02	0.02	0.40
0.56	0.49	0.23	0.66	0.10	0.76	0.02	0.02	0.02	0.40
0.58	0.49	0.23	0.66	0.10	0.76	0.02	0.02	0.02	0.40
0.60	0.49	0.23	0.66	0.10	0.76	0.02	0.02	0.02	0.40
0.62	0.49	0.23	0.66	0.10	0.76	0.02	0.02	0.02	0.40
0.64	0.49	0.23	0.66	0.10	0.76	0.02	0.02	0.02	0.40
0.66	0.49	0.23	0.66	0.10	0.76	0.02	0.02	0.02	0.40
0.68	0.49	0.23	0.66	0.10	0.76	0.02	0.02	0.02	0.40
0.70	0.49	0.23	0.66	0.10	0.76	0.02	0.02	0.02	0.40
0.72	0.49	0.23	0.66	0.10	0.76	0.02	0.02	0.02	0.40
0.74	0.49	0.23	0.66	0.10	0.76	0.02	0.02	0.02	0.40
0.76	0.49	0.23	0.66	0.10	0.76	0.02	0.02	0.02	0.40
0.78	0.49	0.23	0.66	0.10	0.76	0.02	0.02	0.02	0.40
0.80	0.49	0.23	0.66	0.10	0.76	0.02	0.02	0.02	0.40
0.82	0.49	0.23	0.66	0.10	0.76	0.02	0.02	0.02	0.40
0.84	0.49	0.23	0.66	0.10	0.76	0.02	0.02	0.02	0.40
0.86	0.49	0.23	0.66	0.10	0.76	0.02	0.02	0.02	0.40
0.88	0.49	0.23	0.66	0.10	0.76	0.02	0.02	0.02	0.40
0.90	0.49	0.23	0.66	0.10	0.76	0.02	0.02	0.02	0.40
0.92	0.49	0.23	0.66	0.10	0.76	0.02	0.02	0.02	0.40
0.94	0.49	0.23	0.66	0.10	0.76	0.02	0.02	0.02	0.40
0.96	0.49	0.23	0.66	0.10	0.76	0.02	0.02	0.02	0.40
0.98	0.49	0.23	0.66	0.10	0.76	0.02	0.02	0.02	0.40
1.00	0.49	0.23	0.66	0.10	0.76	0.02	0.02	0.02	0.40

a) Published evaluation data on additionality and Return on Investment were used in the model to calculate gross and net impacts, outcomes and outputs.

b) The data was placed in 'look-up tables' so that users could see some of the assumptions driving the model. The citations for each of the sources of data were included in comments.

a) Use of published data

Published evaluation data on output additionality factors and Return on Investment additionality factors were used in the model to calculate net impacts, outcomes and outputs (i.e. these factors were multiplied with gross data to provide net data). Initially, YF 'policy product ranges' denoting programme types were mapped onto national IEF sub-themes (to be discussed in Chapter 5) using the Phase 1 Regeneris evaluation report (YF 2010) 'YF Policy Product Range Evaluations' (see Appendix 1). Within this report, YF additionality factors and Return on Investment (ROI) factors were extracted from a number of YF published evaluation reports and from the PWC report (PWC 2009a; PWC 2009b). These factors were then used within the model to calculate net outputs and ROI for each IEF sub-theme.

b) Use of look-up tables

The data were placed in 'look-up tables' so that users could see some of the assumptions driving the model. The calculations in the model referred back to these look-up tables and the citations for each of the sources of data were included in comments.

3.8 Phase 4: Online workshop

The online workshop was chosen as an innovative approach to elicit views and attitudes towards the use of the decision support tool that had been developed in Phase 3. The design of Phase 4 was highly novel and was influenced by the analysis of the online survey findings in two key ways. Firstly, although the survey was able to capture the views and attitudes towards evaluation across the RDA policy cycle, the workshop aimed to reveal participants' reasons, rationalisations and arguments when faced with having to 'instrumentally' use evaluation evidence, evidence synthesis and a decision support tool for decision making. As it was not possible to undertake research 'in the field' given RDA abolition, a virtual decision environment was created to compare decision making processes with and without the use of the decision support tool. Questions arising from the survey were incorporated into the online workshop. This allowed issues that were not necessarily anticipated, but were of importance to respondents, to be explored more fully. Secondly, analysis of the survey findings highlighted the importance of senior management in shaping an 'evidence

based /evaluation culture'. Therefore, a decision was made to purposefully sample senior policy makers and analysts from the study population.

An online workshop was created to host the virtual decision environment so that individuals could participate at their 'own pace'. This meant that senior, busy and geographically spread officers could be approached to contribute to this research. The method was not only an innovative approach to data collection, but it was practical and cost-effective. A face-to-face workshop was dismissed on grounds of being unfeasible given such parameters.

3.8.1 Background

The online workshop was designed around two decision making scenarios to compare decision making processes without and with the use of the decision support tool (scenario 1 and 2 respectively). The term 'online workshop' was used as an umbrella to describe the range of methodologies that were drawn upon to meet the research aims of Phase 4. Firstly, a virtual decision environment was created to collate quantitative data for the first scenario on the instrumental use of evaluation evidence for decision making. Such a method is most closely linked to *experimental/laboratory approaches*. Secondly, the *decision tool* (developed in Phase 3) was presented to participants to elicit their views and attitudes towards the use of such approaches for the second scenario. Thirdly, participants were asked to self-reflect upon their decision making processes and were asked further questions about their views and attitudes towards the potential role for evaluation evidence and decision support within the regional policy context. Such an approach utilised *online survey* methodology to collate quantitative and qualitative data. Finally, the online workshop was used as a mechanism to host the virtual decision environment and collate the data through the surveys. Within the workshop format I was able to present the scenarios, give the necessary background information and provide an overview of the decision tool's capabilities. Such a methodology is closely linked to *online focus group* methodologies (Tates et al. 2009). Each of these methodologies is discussed in turn below.

For scenario 1, the aim of creating the virtual decision environment was to explore the participants' reasons, rationalisations and arguments when faced with having to 'instrumentally' use evaluation evidence within an investment prioritisation decision

making scenario. Indeed, Papadakis and Barwise (1999, p289) reflect that “one problem with Strategic Decision Making research is that it is rarely possible to observe the process and its characteristics during real time.” The methodology included elements of ‘empirically observing behaviour’ whereby participants were asked to make decisions within the online workshop and submit responses. Such an approach was based upon experimental laboratory design. Bryman (2102, p50) observes that experiments are quite unusual in sociological research but are employed more often within related areas such as social psychology and organisational studies. Initially, it was planned that participants would be allocated to two experimental groups, whereby both groups would be presented with the same scenario but one group would be given the use of the decision support tool to aid decision making and the other group would not. It was anticipated that the decision outcomes could then be analysed to understand the influence of the decision tool.

However, the aim of the research was not to explore the influence of a decision tool on the outcomes of decision making (i.e. what would constitute a ‘better’ decision anyway?), but to explore participants’ views and attitudes towards the instrumental use of evaluation evidence and decision support within the regional policy context. Therefore, rather than a classical experimental design, a virtual decision environment was created (i.e. a contrived setting) whereby all participants were asked to apply presented evidence to a decision making scenario and submit their responses within the online workshop.

The advantage of such a design was that data could be generated avoiding problems of retrospective reporting (such as ex-post rationalisation, memory failure, etc.) and it was possible to control aspects of the workshop in order to meet the aims of the research. The experiment could also be repeated with multiple participants, generating a dataset. However, laboratory design has long been known to have significant limitations when studying ‘Strategic Decision Making’ (SDM) behaviour (Papadakis and Barwise 1999, p289). Even though the setup of the experiment may be ‘realistic’, an experiment is an artificial situation and there will be differences to behaviour in real life impairing validity (Locke 1986). As such, although the data generated from scenario 1 were of stand-alone interest, the underlying rationale for the methodology employed was to evoke and perhaps intensify participants’

reflections towards the instrumental use of evaluation evidence for decision making within the regional policy context.

For scenario 2, the aim of developing the simple decision support tool (Phase 3) was to elicit participants' views and attitudes towards the use of such approaches. By hosting the decision tool on the online workshop, the goal was to give participant's a tangible experience of the tool. However, practically, using the tool required training and support which would be time consuming and likely to lead to technical user issues. Therefore, a decision was made to deliver a presentation of the tool to participants and to ask for feedback, as one might do when testing recommendations within a face-to-face workshop. A short film was made of the presentation so that it could be hosted on the online workshop platform. The advantage of this approach was that all participants would receive standardised information and the presentation would not need to be given 'synchronously' to all participants at the same time.

Participants were asked to self-reflect on their decision making processes and self-completion surveys (questionnaires) were chosen as a tool to capture participant's views and attitudes within the workshop. Closed questions were used to collate standardised, quantitative data and open-ended questions were used to enable non-standardised responses and allow scope for new ideas to be introduced beyond the pre-set questions (Bryman 2012, p246). The advantage of using such a methodology was that the data could be collated within the online workshop setting.

An online workshop was created to host the virtual decision environment, decision tool and surveys. Online workshops are generally associated with e-learning approaches, which are rapidly increasing in demand within training and educational sectors (OLTF 2011). E-Learning has been demonstrated to increase information retention rates and cut down instruction time (JISC 2009). This was essential given that a lot of complex information was to be conveyed and needed to be understood within a short timeframe within the workshop. However, the aim of the online workshop was to *collect* data to meet the research aims, and such an approach was based upon online focus group design. A "focus group is a type of interview where there are several participants in addition to the facilitator" (Bryman 2012, p663). Initially, it was planned that the workshop would be held 'synchronously' so that the workshop was

held in real time to closely reflect a face-to-face approach. Research has demonstrated that there may be little difference in data quantity or quality between synchronous online focus groups compared with parallel data from conventional face-to-face ones (Underhill and Olmstead 2003). However, given that the sample was senior policy makers and analysts, it was not feasible to schedule in a workshop that everyone would be able to (virtually) attend. Therefore, the workshop was held 'asynchronously' with participants able to log in and out of the workshop over a period of two weeks (see Adriaenssens and Cadman 1999 for an example of an asynchronous focus group study conducted via email).

Often the emphasis of focus group methodology is upon the "interaction between the group members and the joint construction of meaning" (Bryman 2012, p712). To capture this, there was also an 'open forum' with prompt questions, enabling participants to post comments for open discussion. It was made clear that all the information given as part of the workshop would be kept confidential, but that the participant's feedback posted on the workshop's forum would be available for all workshop participants to see. The purpose of open forums was to try and capture some level of insight into collaborative decision making processes. The aim was to initiate a debate, to try and capture differences in organisational approaches, and to elicit additional ideas and responses.

Overall, the methodology employed was highly innovative and has not previously been used to investigate the views of an expert stakeholder group involved in RDA regional policy evaluation.

3.8.2 Data collection

A literature review was conducted to inform the development of the content for the online workshop and surveys. The scope for questioning was potentially vast. Research focusing on areas such as decision theory and behavioural research on how decisions are made was found to extend across the disciplines of: psychology (psychological and cognitive perspectives); economics (organisation theory, strategic management, behavioural economics and neuroeconomics); and mathematics (operations research). Approaches to decision making were found to be diverse ranging from rational choice (Savage 1954) to decision making in disorder (Cohen, March and Olsen 1972) with

rational up-to-a-point (Simon 1957) contained within this spectrum, alongside a range of personal belief approaches.

Focusing on the latter two approaches, an exploratory literature review was conducted focusing on the central debates surrounding cognitive processing for decision making and the factors that may bias the process. This included a review of decision maker utility as it has been demonstrated that utility perceptions impact cognitive processing and influence what information is retrieved and how it is evaluated (McCaughey and Bruning 2010). In addition the influence of affect, the experience of feeling or emotion, was reviewed as it has been shown to “influence the manner in which individuals perceive situations, the motivation of decision behaviours, the degree of decision risk tolerance, and the level and type of information recall people exhibit” (McCaughey and Bruning 2010, p10). Finally, the role of ‘*heuristics*’ errors, defined as potential intuitive processing errors, were reviewed as research has found that simplifications, or mental ‘rules of thumb’ and ‘gut feeling responses’ may have implications for the rationality assumptions of evidence based decision making (McCaughey and Bruning 2010). Finally, the decision context of collaboration was reviewed as group decision making has to take account of individual objectives and biases, in addition to the processes of persuasion and opinion influencing (Bazerman 1998). This literature review provided background insight, particularly for interpreting the data. The exploration of such ideas could be investigated more directly in future research.

In terms of shaping the survey content directly, the literature review revealed factors which could potentially affect the decision making process within the scope of the workshop. Such identified factors which needed to be considered in the workshop design included: the decision makers; the decision situation; phrasing the scenarios in terms of a problem or an opportunity; decision criteria; time; outcomes of any decisions/what is deemed a satisfactory outcome; collaborative decision making; and the role of decision support (theories, tools and techniques). Each of these is discussed in turn below.

Every aspect of the workshop and surveys was thought through in detail. In terms of the decision makers, the aim was to recruit officers from senior positions. This was due to the role of senior management in creating an ‘evaluation culture’ as identified from

the survey. Both policy makers and analysts were recruited. Secondly, in terms of the decision situation, the aim was to focus the workshop on investment prioritisation decision making as this was a key part of the policy process where evaluation evidence could have had more influence within the RDAs, as identified from the survey. A decision was made to focus on the recent austerity cuts, so that the research explicitly acknowledged RDA abolition and recent events. Moreover, it was anticipated that through exploring the decision making processes of budget reduction, study findings would be produced that could be transferable to scenarios of investment prioritisation more generally and particularly for ongoing spatial policy. This decision situation influenced the phrasing of the scenarios in terms of a problem or an opportunity and informed the decision criteria.

Thirdly, in terms of time, it was very important that the participants were able to conduct the research at their own pace so that senior, busy officers could be included. However, to ensure that the workshop was actually completed, a timeframe of two weeks (10 business days) was set. Fourthly, in terms of what was deemed as a satisfactory outcome, it was stressed to the participants that there was no right and wrong answer; instead, it was the thought processes they went through when undertaking the workshop which were of interest for this research. Fifthly, in terms of collaborative decision making, initially it was planned to segment the research participants into smaller groups and to structure the scenarios and assignments so that collaboration could also be investigated. However, it was felt that this would lead to a certain level of frustration in completing the workshop, which may have led to higher attrition. Additionally, it was felt that the process would not be able to simulate the context of collaborative decision making closely enough for the research to be valid. Therefore, the open forum was used to capture the discussion of ideas. Finally, in terms of the role of decision support, the workshop aimed to specifically investigate the attitudes and views of participants towards a presented decision support tool.

Data were collected through bespoke surveys embedded within the online workshop web-pages (Appendix 7). As a (previously) practising RDA evaluation officer I had some prior understanding of what elements of content and structure might be used for the workshop and surveys. This background knowledge of policy and practice, and the context of policy change, enabled me to devise data collection tools. The surveys were

designed using a simple format including: multiple choice asking for one option; self-assessment items measured on a 5-point/3-point Likert type scale; dichotomous answers like “Yes” and “No”; and open-ended questions, qualitative questions. Some questions in the survey had an open-ended “Other (specify)” option. A choice of “Not applicable” (NA) was included when necessary.

The workshop format

The workshop was web-based. Based upon findings from a scoping interview with an e-learning specialist, the workshop was designed and hosted on a Ning website (www.ning.com), a platform for creating social websites frequently used for the administration of online educational and training courses (Clark 2011. pers. comm.) The online workshop aimed to follow the conventional approach of a face-to-face workshop as much as possible. The language used throughout the workshop was kept simple, to the point and ‘friendly’. For instance, participants were thanked for their time at the beginning and end of the workshop. The format of the online workshop is described in detail below.

Home web-page

A screenshot of the workshop homepage is presented in Figure 7.

Figure 7: Screenshot of the workshop homepage

WORKSHOP: DECISION MAKING
Investment Prioritisation in UK Regional Policy

WELCOME
Welcome to the workshop!

- Other workshop participants can be viewed to the right and their details can be viewed [HERE](#).

Please note:

- You can log in and out of the workshop when you like using your email and password.
- You may wish to have headphones to listen to the videos in the workshop as they are narrated.
- When submitting the surveys, please ensure you click 'done' to submit the survey results and ensure you see a thank you page before continuing with the workshop.
- If you have any problems, please email: Jessica.K.Baxendale@student.she.ac.uk or call 07590457394. During office hours, immediate assistance should also be available via the chat function at the bottom of the page.

The workshop is set out as follows:

- Introduction: An overview of the workshop
- Scenario 1: Investment prioritisation exercise
- Scenario 2: Investment prioritisation using the decision tool
- Phone call: An overview of what will be discussed in the follow-up telephone interview
- There are forums for scenario 1 and scenario 2 so that all workshop participants can discuss the issues, comment and ask questions. Please contribute to these forums.
- The latest posts from the forums will appear below on the home page

THANK YOU

FORUM

Scenario 1: How did you weigh up the different pieces of information to come to a decision? 1 Reply
Started by Jessica Baxendale in Scenario 2: Investment Prioritisation using the Decision Tool. Last reply by Robin Beveridge 19 hours ago.

Scenario 2: How do you think the accuracy of monitoring data, socio-economic data, and evaluation data can be improved? 2 Replies
Started by Jessica Baxendale in Scenario 2: Investment Prioritisation using the Decision Tool. Last reply by Robin Beveridge 19 hours ago.

Scenario 1: How did you weigh up the different pieces of information to come to a decision? 2 Replies
Started by Jessica Baxendale in Scenario 1: Investment Prioritisation. Last reply by Robin Beveridge yesterday.

[+ Add a Discussion](#)

MEMBERS

[Edit](#)

[View All](#)

Members Online (3)

- Jessica Baxendale
- Paul Mooney
- Sarah McMillan

JESSICA BAXENDALE
Sign Out
inbox (1 new)
Friends - none
Settings

Centre for Regional Economic and Social Research
Sheffield Hallam University

[Edit](#)

- b) The homepage orientated participants to the following pages:
- Introduction page
 - Scenario 1
 - Scenario 2
 - Follow up phone-call

- d) An open forum was used to enable open discussion and debate.

- a) Participants were able to log in and out of the workshop over a 2 week period. This flexibility enabled respondents to be targeted from senior positions.

- c) Photos of participants were used to create the sense of a face-to-face workshop.

- e) A chat function provided technical support. Participants were able to see who else was 'online.'

a) Log-in

An online workshop was created to host the virtual decision environment so that individuals could participate at their 'own pace'. To keep it flexible, participants were able to log in and out of the workshop when they liked using their email and password. When participants had completed the workshop they were informed they were still able to log into the workshop to participate in the open forum.

b) The workshop webpages

The homepage was used primarily to orientate participants around the workshop webpages, which flowed sequentially:

- Introduction page: giving an overview of the workshop
- Scenario 1: Investment prioritisation exercise
- Scenario 2: Investment prioritisation exercise using the decision tool
- Follow up: An overview of a follow-up telephone interview

There were also 'open forums' for scenario 1 and 2 so that all workshop participants could discuss the issues raised and ask questions openly. The latest posts from the forum appeared on the homepage. Participants were also able to see who else was 'online'. Respondents could return to the homepage from any web-page within the workshop by clicking on the 'home' tab. Further detail on the other aspects of the homepage and design of the workshop will be discussed below.

Introduction web-page

A screenshot of the introduction is presented in Appendix 8. On the introduction page participants were thanked for taking part. The purpose of the workshop and the practicalities for completing the workshop were detailed.

Scenario 1 web-page

Screenshots of the scenario 1 web-page are presented in Figure 8 and Figure 9. Participants were advised to spend approximately 30 minutes on scenario 1.

Figure 8: Screenshot of the workshop scenario 1 (top of the screen)

An introduction to scenario 1 was given. Participants were asked to prioritise investments (i.e. rank projects) without the use of a decision tool.

Scenario 1

Introduction

- You are charged with advising the Board of a new Development Agency on its project spending priorities for 2012-13
- The evaluation team in the Development Agency has drawn together a summary of 10 economic development projects, all of which are eligible for funding
- The Government has announced reductions in the Agency's budget and it will now only have £5 million to spend in the forthcoming year instead of £10 million
- Your task is to prioritise investment expenditure based on the available evidence.

Here is the summary of the 10 economic development projects

Scenario 1: Projects

Project name	Project description	Local Authority	Forecast investment 2012-2013	Total gov. in investment 2012-13	Jobs created	Businesses created	Businesses supported	People assisted to gain employment	People assisted in skills	CO2 reduced (tonnes)	Political/pragmatic considerations
1. Digital House	Building purchased to develop into high quality office accommodation for digital companies	Sub-regional	£400,000	£8,200,000	100	0	0	0	0	0	Took 5 years to acquire working with the local authority. Currently an eyesore in the centre of the town. A lot of negative publicity in the local press about how long it has taken to begin development.
2. Regional events	Funding for regional events to raise the profile of the region and attract inward trade, investment and visitor spend	Pan-regional	£600,000	£8,460,000	0	0	0	0	0	0	It is difficult to capture the impact of regional events. However they are high profile events.
3. Above market regeneration	Regeneration of town market complex into a theatre, cafe and town council offices	District C	£300,000	£4,230,000	5	5	20	0	0	0	Most of this project has been carried out. This remaining budget is for completion and doing an evaluation to look at the impact of the investment.
4. Charity support	Support advice and funding for the voluntary and	Pan-regional	£2,500,000	£24,500,000	0	2	70	0	0	0	The charity sector is going to become more important going forward as public services are

Projects

- Take 5 minutes to read through the table of information. You may wish to view/print out the spreadsheet [HERE](#).
- The economic development projects listed are based on actual previous projects, but the data is dummy data.

Data on 10 hypothetical RDA projects were given. This spreadsheet was embedded into the webpage.

Figure 9: Screenshot of the workshop scenario 1 (bottom of the screen)

Record your decisions here and answer the following questions:

- You may wish to use a calculator or you can use the functions in the spreadsheet [HERE](#).
- Please ensure you click 'done' to submit the survey results at the end and ensure you see a 'thank you' page before continuing with the workshop.

Scenario 1

*1. What is your full name?

*2. Rank the projects from 1 to 10 (1 = most favourable project to receive investment, 10 = least favourable project to receive investment).

	1	2	3	4	5	6	7	8	9	10
1. Digital House	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2. Regional events	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3. Atown market regeneration	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4. Charity support	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5. Atown railway station	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Create your free online surveys with SurveyMonkey, the world's leading questionnaire tool

To finish:

- Please ensure you have clicked 'done' to submit the survey results above and ensure you have seen a 'thank you' page before continuing with the workshop.
- Answer the questions (or start a new topic) on the public forum for scenario 1 [HERE](#) to discuss the issues with fellow workshop participants.
- Next on to scenario 2 [HERE](#)

Data on participant's decision making processes were collated through an online survey.

The survey was embedded into the workshop webpage.

Participants were asked to imagine that they were charged with advising the Board of a new Development Agency on its project spending priorities. The evaluation team in the Development Agency had drawn together a summary of 10 economic development projects, all of which were eligible for funding. The government had announced reductions in the Agency's budget and it would now have £5 million to spend in the forthcoming year instead of £10 million. The task was to prioritise investment expenditure based on the available evidence.

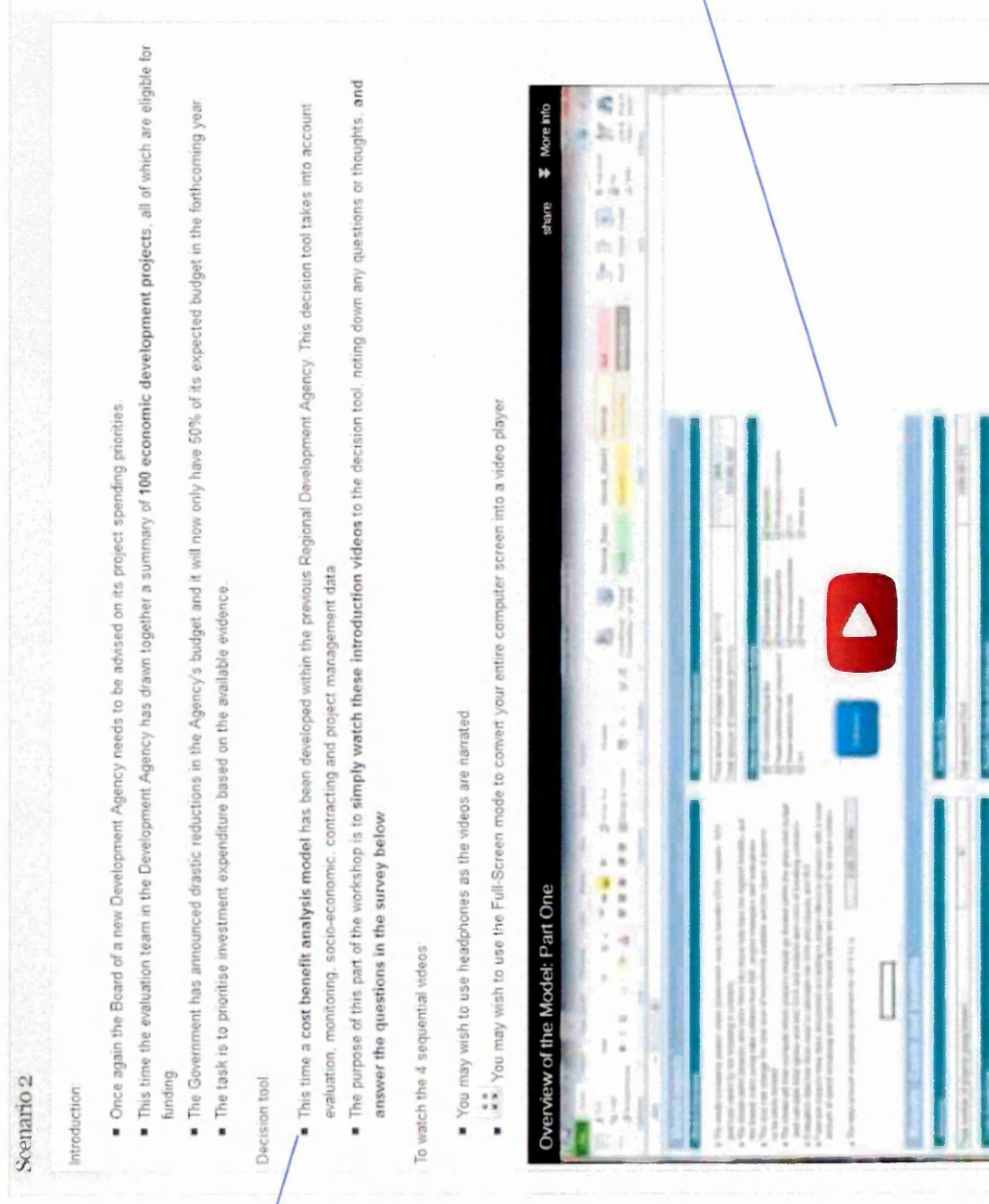
As shown in Figure 8, data on 10 economic development projects were given. The projects were based on previous projects that may have been feasibly funded by an RDA, but the data given was dummy data. A Microsoft Excel spreadsheet containing the data was transformed into a web-page spreadsheet using google docs (<https://docs.google.com>) and then embedded within the workshop using the functionality in Ning. A link was also provided so that participants were able to download a Microsoft Excel spreadsheet to view or print out. The data presented to participants for scenario 1 is provided in Appendix 9. As shown in Figure 9, data on participants' decision making processes was collected through a bespoke survey. The survey was web-based and hosted through the SurveyMonkey website (www.SurveyMonkey.com). The survey was embedded within the workshop using the functionality in Ning, meaning that participants were not re-directed to another website and instead stayed within the online workshop web-page.

When scenario 1 was completed, participants were directed to scenario 2.

Scenario 2 web-page

Screenshots of the scenario 2 web-page are presented in Figure 10 and Figure 11. Participants were advised to spend approximately 20 minutes on scenario 2.

Figure 10: Screenshot of the workshop scenario 2 (top of the screen)



An introduction to scenario 2 was given. Participants were asked to consider prioritising investments (i.e. ranking projects) with the use of a decision tool.

Participants were asked to simply watch an introduction video to the decision support tool, noting down any questions or thoughts. The video was embedded into the workshop webpage.

Figure 11: Screenshot of the workshop scenario 2 (bottom of the screen)

Record your answers to the following questions here

- Please ensure you click 'done' to submit the survey results at the end and ensure you see a 'thank you' page before continuing with the workshop.

Scenario 2

*1. What is your full name?

*2. How do you think this scenario compares to scenario 1?

*3. Would you use such a tool for decision making?

Create your free online surveys with SurveyMonkey, the world's leading questionnaire tool.

To finish:

- Please ensure you have clicked 'done' to submit the survey results above and ensure you have seen a 'thank you' page before continuing with the workshop.
- Answer the questions for scenario 2 (or start a new topic) on the public forum [HERE](#) to discuss the issues with fellow workshop participants.
- Your thoughts on the model will also be discussed in the telephone interview.
- View what the telephone interview will cover [HERE](#).

Data on participant's perspectives on the decision support tool were collated through an online survey.

The survey was embedded into the workshop webpage.

Once again, participants were asked to imagine that the Board of a new Development Agency needed to be advised on its project spending priorities. Again, the Agency's budget was to be reduced by 50 per cent and the task was to prioritise investment expenditure based on the available evidence. However, in scenario 2 one hundred economic development projects were presented (compared to 10 in scenario 1). In scenario 2 an overview of a decision support tool was presented to participants via four sequential online videos including: an introduction; the data; the inputs; and the reporting of the decision tool. Participants were asked to simply watch these introduction videos to the decision tool, noting down any questions or thoughts, and answer the questions in the following survey.

Based upon the scoping interview with the e-learning specialist, videos were recorded using Camtasia Studio 7 software⁴⁵ (www.techsmith.com/Camtasia). To record videos, online tutorials for the software were followed, an appropriate script was written and the sound recorded and edited. Then the software was used to record PC screen movements in time with the voice recording. This enabled the decision tool to be demonstrated visually to participants. The videos were uploaded to YouTube (www.youtube.com) and the correct security settings were enabled so that the videos were not publically available. The YouTube links were then embedded within the workshop using the functionality in Ning.

As shown in Figure 11, participants' perspectives on the decision support tool were captured through a bespoke survey. The survey was, once again, web-based and hosted through the SurveyMonkey website (www.SurveyMonkey.com).

When scenario 2 was completed, participants were directed to the next web-page about the follow-up telephone interviews.

Follow-up telephone interviews web-page

A screenshot of the web-page giving detail on the follow-up telephone interviews is presented in Appendix 8. The follow-on telephone interviews were booked in advance with participants and were primarily used as a means of ensuring that participants completed the workshop by the agreed date of the interview. 18 participants had a

⁴⁵ Camtasia is used for screen recording and video editing.

follow-on interview. Interviews were 30 minutes long, semi-structured and were conducted over the phone. One participant had their interview conducted over Skype for efficiency reasons as they were abroad. The topic guide for the interviews is included in Appendix 10. The interviews were recorded, but the data was not transcribed. Given the volume and quality of data collected from the online survey and workshop, the interview data has not been included as part of this thesis. Such data could, however, form the basis of future research work.

b) Photographs

An experimental component of the workshop was to ask participants to upload a photograph of them as part of the initial log-in process. This is shown in Figure 7. The idea was to try and recreate a sense of a face-to-face workshop. It was also felt that, as many officers involved in RDA evaluation knew each other professionally, having photographs and the names of workshop members displayed may have created a sense of openness, rather than a sense of secrecy if members were anonymised. As many of the participants held senior roles, their visual presence in the workshop added a sense of legitimacy and thus may have encouraged greater participant involvement and lower attrition rates. However, not every participant chose to upload a photograph and it did add another hurdle to the log-in process, which was a risk.

c) The open forum

Another experimental component of the workshop was to develop an 'open forum' using the functionality in Ning. It was made clear that all the information given as part of the workshop would be kept confidential, but that the participant's feedback posted on the workshop's forum would be available for all workshop participants to see. The purpose of open forums was to try and capture some level of insight into collaborative decision making processes. The aim was to initiate a debate, to try and capture differences in organisational approaches, and to elicit additional ideas and responses. Links to the forum were clearly marked at the bottom of each scenario web-page and the latest forum posts were presented on the homepage. Prompt questions were used to initiate debate, but participants were encouraged to start new discussions and to ask questions openly to the group. Participants were also able to see who else was 'online' to enable synchronous debate. However, as this was not a mandatory part of the workshop, many participants chose not to make comments on open forums.

The questions posed in the forums (and the corresponding response rate) were as follows:

- Scenario 1: In real life, how would you have decided which projects should go forward? – 5 respondents
- Scenario 1: How did you weigh up the different pieces of information to come to a decision? – 4 respondents
- Scenario 2: How do you think the accuracy of monitoring data, socio-economic data, and evaluation data can be improved? – 3 respondents
- Scenario 2: Could such models be applied to decision making in the future? – 2 respondents
- Scenario 2: How could the model be improved? – 0 respondents

This is a relatively low response rate and there was not enough data collected to compare organisational approaches to evaluation and decision making. Nevertheless, the responses given were thoughtful and showed an interesting level of debate between the participants.

d) Technical support

Technical support was a key issue to be considered to reduce the risk of attrition or missing data. Participants were given a number of options. They were able to email or phone for technical assistance. They were also able to use a private 'chat' function to ask for immediate, visual support to be provided. This was an experimental component of the workshop. It was felt that participants may be more inclined to ask for assistance through this indirect method than through phoning or having to explain the issue in a more formal email. However, providing a chat function meant that it was necessary to be online to offer support continuously for the two-week period that the workshop was administered. To manage expectations it was made clear that the chat function would operate during office hours only.

Pilot testing

To reduce the need for technical support, the workshop was thoroughly pilot tested. It was very important for the workshop to be easy to use and for the videos and decision tool to look professional to ensure participant engagement, to maximise the response

rate, and to reduce the risk of attrition. The use of Ning, Camtasia, SurveyMonkey and YouTube not only meant that the workshop looked professional, but it enabled the workshop to flow sequentially without participants leaving the workshop web-pages.

The workshop was extensively piloted prior to use in order to iron out any user issues. Piloting was undertaken with the PhD supervisory team. In addition, the full workshop was 'road tested' by other PhD students and then critically presented and discussed within a PhD Forum seminar. Finally, scenario 1 and the concept of the workshop were tested formally with Sheffield Hallam University Masters students within group work and feedback during an evaluation seminar. The findings from these tests were incorporated into the final version of the workshop. In particular, it was ensured that any documentation to be downloaded was saved in earlier versions of Microsoft Excel/Word to solve the problem of software incompatibility.

Administering the workshop

All efforts were made to maximise the response rate. Gatekeepers (managers) were contacted directly in BIS, emda and ONE initially and the purpose of the research explained in detail. The gatekeepers were then asked to provide a sample of ten participants and their contact details for each organisation. Potential participants were then contacted via email or through the messaging service of the business-related social networking site, Linked In (www.linkedin.com). Contact was made via an introductory email stating that the individual had been recommended for the research by the gatekeeper, information was provided to build trust such as researcher background information, and confidentiality was emphasised.

The email included a link to an online participant information sheet and consent form (see Appendix 11). As part of the consent form, participants were asked to book in a time for their telephone interview using an online booking facility (www.doodle.com) and to provide a contact phone number. The aim was to capture all the information required in one contact to reduce the chance of attrition. Other standardised communication included:⁴⁶

⁴⁶ See Appendix 12 for all standardised communication.

- One week prior to the workshop a reminder and log in details were sent with instructions of how to navigate around the online workshop.
- On the 25th June 2012, an email was sent out stating that the workshop had gone 'live'.
- When participants completed the workshop (i.e. submitted their survey responses) a thank you email was sent.
- When participants completed the telephone interview a thank you email was sent.

A communication tracker document was kept. The workshop officially 'closed' on the 6th July 2012. Participants had a scheduled time for their telephone interview to be conducted when they had completed the workshop (which ensured the workshop was completed on time).

Sequential snowball sampling was also adopted to recruit additional participants when an opportunity presented itself and participants proposed other participants who had the experience or characteristics relevant to the research. Therefore three more participants joined the workshop after the 25th June start date.

To reduce attrition, it was stressed to the participants that there were no right and wrong answers and that it was their thought processes, and the usefulness of evidence to assist those decisions, that was the focus of the research. It was also emphasised to the participants that it was their decision about how long they wanted to take to complete the workshop. The minimum amount of time it should have taken was one hour to watch the videos and complete the tasks quickly.

Ethical considerations

The proforma that was completed to gain ethical approval for the research is discussed in section 3.6. Informed consent was gained for the online workshop and telephone interviews. Participants were provided with an online 'participant information sheet' and consent form prior to the workshop and interviews. One consideration with obtaining consent through an online format was whether or not a signature would need to be provided. However, it was clearly stated in the participant information sheet that the online consent form was understood to mean that informed consent to

participate in the study had been given. This online consent form was reviewed by the Head of Research Ethics at Sheffield Hallam University.⁴⁷ There was also a link to a Word document version of the participant information sheet and consent form on the introduction web-page of the online workshop for participants to download and keep for future reference (see Appendix 11 for a copy of the form). Once again, the confidentiality of information was maintained and the anonymity of participants respected. The research did not cause harm to the participants, involve sensitive topics, or involve vulnerable groups (Bryman 2012, p146-7).

Participants

Participants were recruited to the workshop based upon their role, skills and experience. Respondents were targeted from senior positions from the Regional Development Agencies (RDAs) and central government (see Appendix 13 for the list of participants). A total of 19 senior policy makers were recruited from three organisations including: former East Midlands Development Agency (EMDA) officers; former One North East (ONE) officers; and central government officers from the Department of Business, Innovation and Skills (BIS); and the Department for Communities and Local Government (CLG). A discussion of the characteristics of the survey respondents is detailed in Chapter 8 (section 8.2). Although nineteen participants may be considered a small sample size, the respondents were an expert stakeholder group with background knowledge and insights which were of direct relevance to the research topic. Participants were asked to devote at least 1.5 hours⁴⁸ to participate in the workshop, which is a substantial amount of time. In addition, Adriaenssens and Cadman (1999) suggest that, for asynchronous focus group studies, small groups of participants are most effective due to potential research management issues.

Reflections on the data collection

The online workshop did not take a conventional approach, and a range of methodologies were drawn upon and 'mixed' to meet the research aims of Phase 4. A

⁴⁷ Head of Research Ethics, Sheffield Hallam University. Email communication 14.05.2012.

⁴⁸ One hour for the workshop and thirty minutes for the telephone interview.

more conventional approach would have been to either use formal participatory observation techniques (Bryman 2012, p714) during actual decision making processes within the RDAs or to have used a laboratory approach more closely aligned with classical experimental research, such as behavioural economics methods (Foote, Goette and Meier 2009). However, given the timing of the research and the aims of Phase 4, the online workshop provided a platform to convey a lot of information in a short amount of time and to generate qualitative and quantitative data. As expected, the pool of participants for the online workshop is not representative; therefore only qualitative aspects of the results are expected to be reliable. However, the quantitative results do frame the discussion and indicate the underlying trends in the data. It is the qualitative data that are most illuminating and which underpin the key findings. The neutral online venue and the subsequent dynamic of communication in response to the decision tool led to thoughtful responses and encouraged participants to open up and contribute. Therefore the workshop has provided a rich source of quantitative and qualitative data. It could also be argued that the scenarios presented avoided problems of retrospective reporting and memory failure.

Once again, an issue relating to validity was my known identity as a previous RDA evaluation officer and my known 'authorship' of the decision tool presented in the workshop, which may have influenced response to some questions. Overall a key limitation was the artificiality of the experiment for scenario 1. In particular participants' approaches to decision making may be different within the workshop compared to reality and the cognitive activity of weighting different sources of information is likely to be an unconscious thought process usually, which may be conducted differently when undertaken 'consciously'. Therefore a key limitation is the self-reported levels of evidence use, which may not be accurate. Response bias is likely, given firstly the phrasing of some of the quantitative questions (in particular the predetermined series of statements) which may have led to issues of evocation and intensification; secondly, the respondent's relationship to the researcher and to evaluation team members; thirdly, the timing of the survey during RDA abolition meaning that stakeholders had highly vested interests; and finally, the potential desires of participants to provide 'textbook responses'.

There were further challenges in conducting this research. Firstly, there were issues of digital access and digital literacy that may have limited who could be included in the research. There was a need for participants to have a degree of technical competence (particularly in using Microsoft Excel) and technical support was needed. There were also challenges to building rapport with participants in comparison to a face-to-face workshop. The forum, in particular, presented challenges in that the open discussion board was beyond researcher control. Secondly, a major challenge was the time needed to complete the workshop. It was essential to keep the workshop short to increase participation rates; however, participants fed back that it was an unrealistic expectation to complete the workshop in one hour as suggested. It is probable that this mostly explains why five people logged into the workshop without starting it and one participant only contributed to open forums (see Table 8 below). Indeed, most participants who completed the workshop reported that they took longer than one hour and took the time to give detailed responses. This leads on to the final point. A major issue of this type of research is the artificiality of the situation. However, while the study may have lacked ‘mundane realism’ it nonetheless benefited from ‘experimental realism’ (Aronson and Carlsmith 1968) meaning that the participants became immersed in the scenarios posed and took the workshop seriously.

As demonstrated by Table 8, once participants began the workshop and completed scenario 1, there was no attrition. In addition, there was only one major error in collating the data whereby one participant was unsuccessful in submitting their completed scenario 1 due to a technical issue caused by internet connection failure. This is quite an achievement given the experimental and complex nature of delivering the workshop, including the range of software used and research management needed.

Table 8: Administering the workshop

	n
Invitations sent	38
Consent form completed	23
Logged into the workshop	23
Completed scenario 1 survey	17*
Completed scenario 2 survey	18
Completed the online forums only	1
Follow-on telephone interview	18

* Technical error in collecting data for 1 participant

3.8.3 Data analysis

Data analysis and interpretation is presented in Chapter 9. The data analysis conducted for the workshop reflects that conducted for the online survey (see section 3.6.2).

Reflections on the data analysis

To study the role of organisational context in more detail, ideally it would have been interesting to conduct thematic analysis within and across contexts (organisations/departments) and cases (individuals) (Bryman 2012, p417). However, given the large volume of data and the small sub-group sample sizes, such analysis was deemed unfeasible given the scope of this research. This could form the basis of future work.

3.9 Conclusion

This chapter has described the approach, methodology and procedures followed for the empirical research. Overall, the context of this research (RDA abolition) created a situation whereby the study population became hard to reach. To overcome such barriers, the development of the online workshop was a highly innovative approach to data collection, generating lessons from the research management process.

It was found that developing and managing an online workshop is a time intensive process. The use of 'Ning', 'Camtasia', 'GoogleDocs' and 'SurveyMonkey' software to develop the workshop was effective and produced a professional end result, but required researcher training. The importance of using gatekeepers to help recruit participants, keeping the total number of participants to a manageable level and using standardised communication to reduce the duplication of effort was revealed.

Given that such a workshop required participants to have a certain level of digital literacy, and that it was fundamentally important for data to be collected without error, the role of technical support and the need to thoroughly pilot test the workshop was emphasised. In this instance, the use of a 'chat function' to provide technical support was rarely used by participants and was time intensive to manage. It is likely that clear documentation and instructions outlining how to use the workshop were important for reducing the need for such support.

A need was revealed to identify issues which may lead to participant attrition. In recruiting participants for the workshop, there was a benefit in capturing all the information required in one contact (i.e. booking in a time for the telephone interview at the same time as gaining informed consent). It was also essential that all of the distinct elements of the workshop (surveys, online videos, etc.) were embedded within the workshop web-pages and that the workshop flowed sequentially. Consideration was needed to balance the needs of the research (in terms of the workshop content), against the time it took participants to complete the workshop in reality. Finally, having a telephone interview booked in at the end of the workshop was an effective approach to ensure that participants finished the workshop by the stated deadline.

It was found that to replicate more closely a face-to-face focus group and to generate some level of group dialogue, the use of photographs and open forums could have been more embedded as a mandatory part of the workshop. However, the effect this would have in terms of the time needed to complete the workshop and attrition would need careful consideration and this may have led to issues of anonymity and confidentiality. Overall, running the workshop online and 'asynchronously' was effective at enabling busy, senior and geographically spread individuals to participate in the workshop. It was also cost effective. These original research management insights extend the current online research methods literature.

The next part of the thesis presents the research findings. The following three chapters provide a comparative review of the academic and policy literature to address the first research question: what are the epistemological and applicability challenges of extending an Evidence Based Medicine approach to regional policy evaluation? The overarching EBPM debates will be explored focusing on a comparative study of health policy and EBM (Chapter 4) and regional policy and impact evaluation (Chapter 5) to be analysed and interpreted in Chapter 6.

Chapter 4

Health Policy & Evidence Based Medicine

4.1 Introduction

This chapter draws upon methodological guidelines for economic evaluation and the academic and policy literature to contribute towards addressing the first research question: what are the implications of extending an Evidence Based Medicine (EBM) approach to regional policy evaluation?

This chapter provides the foundation for the analysis conducted in Chapter 6 whereby EBM approaches are used as a yardstick against which wider social policy is assessed and parallels are drawn between the practices of EBM and EBRPM. Focusing on the case study of health policy and EBM, the analysis in this chapter is structured by the three central theoretical questions found in the EBPM literature: what kinds of evidence are used and the role of research credibility ('what'); how is evidence incorporated into policy making ('how'); and what are the other factors besides evidence which affect the way policy is made ('other factors'). The following will therefore be discussed: the dominance ascribed to experimental and systematic review research and clinical expertise evidence; the role of the National Institute for Health and Care Excellence (NICE) in incorporating evidence into the policy making process; and the role of other factors besides evidence such as patient preferences. To finish, the challenges of extending an EBM approach to the wider (public) health and social EBPM agenda will be explored. The structure of this chapter is mirrored in the succeeding chapter which is focused on regional policy and impact evaluation.

4.1.1 Theoretical background

This chapter predominantly focuses on debates around evidence types and the role of research credibility. Therefore the theoretical background for this chapter is provided in Chapter 2 (section 2.3).

4.1.2 Scope: Health policy in England

Health policy in England was chosen as a case study to explore the development of EBM. The discussion that follows is relatively simplified and concise. Health policy

development including theoretical models of health care, the levels of delivery and the changing aims, institutions and participants of health policy as well as developments within EBM and extensions of EBM are complex, highly debated topics worthy of detailed consideration in their own right. The purpose of this chapter is primarily to discuss the EBM approach as a means of presenting one end of the EBPM spectrum and then in Chapter 6 to draw out the implications of extending this approach to more complex policy arenas beyond medicine. To demonstrate the EBM approach in England, the Health Technology Assessment (HTA) model of the National Institute for Health and Care Excellence (NICE) is explored. HTA is a process employed by NICE “to examine the safety, clinical efficacy and effectiveness, cost, cost-effectiveness, organisational implications, social consequences, legal and ethical considerations of the application of a health technology, usually a drug, medical device or clinical/surgical procedure” (NICE 2013b).

This focus on medical treatments and technologies through the HTA process of NICE is clearly distinct from studying the role of the evaluation of health care delivery, which may include the evaluation of business change, training and knowledge management (NHS 2014). Beyond HTA, NICE clinical/social care guidelines are not officially mandatory (i.e. there is no legal requirement to follow them), although they are often used as the basis of regulation handbooks and inspection. Of interest when considering health care delivery more widely, is that there is a strong commitment to evidence based decision making and to evaluation within NHS strategic documentation (NHS England 2015). Indeed, the use of pilot schemes to test new innovations in care models and the greater use of observational studies and RCT’s embedded within routine general practice and clinical care are outlined in the NHS ‘Five Year Forward View Strategy’ (NHS 2014). However, the challenges of extending an EBM approach to wider health and social policy are briefly examined in the final part of this chapter.

4.2 Evidence Based Medicine (EBM)

4.2.1 A biomedical model of health

At first sight, and in its most simplistic form, EBM with a biomedical model of health at its core could appear to provide an ideal framework for standardised health care. Attention is drawn to a seminal paper by Engel (1977) in which biomedicine was

identified (and critiqued) as the dominant paradigm of defining disease within western health care systems, with molecular biology the underpinning scientific discipline. Emphasis was placed upon the physical nature of disease and on the treatment of *individuals*. Jones (1994) explains that within a biomedical model of health care, health is predominantly viewed as the 'absence of disease' and the goal of health services is to enable individuals to reach a level of 'functional fitness'. Pascall (2007, p419) notes that "a biomedical model of health was the dominant model at the beginning of the NHS." However, this approach focused on the physical nature of disease (over psychological and social factors) has received criticism for its 'reductionist' framework (Engel 1977).

Gerber, Hentzelt and Lauterbach (2007) reflect that changing concepts or definitions of health have engaged with health policy (macro level), clinical practice (micro level), and research (meso level). Focusing on the latter point, Somekh et al. (2005) argue that health research literature has tended to be dominated by the single discipline of medicine and thus by the natural sciences. Pascall (2007, p419), on the other hand, notes that "medicine's role in health has been openly challenged by social science." Further challenges have arisen from: the development of 'patient consumerism' whereby patients are more involved in decisions about their own health care; the changing balance between the roles of doctors and other professionals; the development of social and environmental theoretical models of health care delivery; and a focus on preventative, as well as diagnostic or therapeutic, interventions (Pascall 2007). Despite this, EBM has gained influence within the health sector and the agenda has recently changed and broadened to extend the EBM approach to wider health and social care policy (Great Britain 2012).

4.3 Evidence types and the role of research credibility

The EBM literature has tended to focus on the micro level (i.e. on encounters between patients and doctors) and explores questions of evidence production and on problematising what is considered relevant evidence to inform policy and practice. Haynes, Devereaux and Guyatt (2002) note that, traditionally, individual doctors' expertise formed the basis of clinical practice. Cochrane's seminal text 'Effectiveness and Efficiency' (1972) marked a shift in thinking, urging the greater use of *external* research evidence within clinical decision making to improve the quality and safety of

health care. Although this signified an early call for EBM, the movement described as a “new paradigm for medical practice” (EBM Working Group 1992, p2421) formally began in the early 1990s. It was spurred on by both an increase in the publication of medical literature and improvements in the accessibility of such research, leading to an influx of information requiring critical appraisal by decision makers. EBM was described by Sackett et al. (1996) as aiming to “promote an explicit and rational process for clinical decisions...emphasizing the importance of incorporating the best research findings into clinical care.” Haynes, Devereaux and Guyatt (2002, p2) note that in turn health research was expected to “meet the dual requirements of being both scientifically valid and ready for clinical application.”

Interestingly, early formulations of EBM de-emphasised doctors’ professional experiences and skills. This is clearly demonstrated in one of the evidence hierarchies reproduced by Guyatt et al. (Table 9) to assist the *critical* appraisal of literature. The evidence was classified into levels based upon study design, and traditional forms of evidence such as observational studies, physiologic studies, patient testimonials, and case reports alongside ‘unsystematic clinical observations’ were regarded as weaker evidence.

Table 9: A hierarchy of strength of evidence for prevention and treatment decisions

-
- N of 1 randomised controlled trial
 - Systematic reviews of randomised trials
 - Single randomised trial
 - Systematic review of observational studies addressing patient-important outcomes
 - Single observational study addressing patient-important outcomes
 - Physiologic studies (studies of blood pressure, cardiac output, exercise capacity, bone density, and so forth)
 - Unsystematic clinical observations
-

Source: Guyatt et al. 2008, p11

In their paper describing the evolution of the EBM approach, Satterfield et al. (2009, p371) note that after “critical exchanges within the medical community, EBM was more explicitly defined as ‘the conscientious and judicious use of current best evidence from clinical care research in the management of individual patients’ (citing Sackett et

al. 1996, p71).” This more mature concept of EBM acknowledged the importance of *clinical expertise* evidence. An updated model advocated the value of doctors’ professional judgement directing EBM decision making (Haynes, Devereaux and Guyatt 2002). In their recent literature review, Rutter, Hawkins and Parkhurst (2013) discuss that renewed emphasis and importance has been placed on experiential and expert (tacit) knowledge within health.

However, there remains a focus on promoting *rational* decision making and evidence hierarchies for appraising the quality of external evidence remain a central component of the EBM approach. Nutley, Powell and Davies (2013, p10) reviewed the EBM literature and conclude that evidence “hierarchies have much in common.” Often randomised controlled experiments with clearly defined controls (RCTs) pioneered within medicine (Medical Research Council 1948) are advocated to be the ‘gold standard’ approach for study design (Grossman and Mackenzie 2005). In particular, the strongest RCT design for therapeutic interventions are considered to be triple-blind, placebo-controlled trials with allocation concealment and complete follow-up involving a homogeneous patient population and medical condition (Quick et al. 2013). The status accorded to research synthesis approaches can vary across hierarchies, but it is widely accepted that single study findings are potentially misleading and open to bias (Chalmers 2007). In particular, ‘systematic reviews’ (Cochrane 1979) are critical to the EBM model, as highlighted in this extended quotation from the Centre for Reviews and Dissemination (2008, pv):

Systematic reviews aim to identify, evaluate and summarise the findings of all relevant individual studies, thereby making the available evidence more accessible to decision makers. When appropriate, it is argued that combining the results of several studies may give a more reliable and precise estimate of an intervention’s effectiveness than one study alone.

Thus importance is also ascribed to statistical meta-analyses (Glass 1976). Qualitative evidence generally has a lower status than quantitative evidence as knowledge, with case study reports usually at the bottom of the hierarchy because of the lack of a control group and the biases inherent in observation and reporting (Guyatt et al. 2008).

An update to traditional hierarchies was put forward by the Grading of Recommendations Assessment, Development and Evaluation (GRADE) Working Group (Atkins et al. 2004). Bagshaw and Bellomo (2008) note that in addition to aspects of (internal) validity highlighted by hierarchies focused on study design (i.e. randomised trial versus observational studies), the GRADE system acknowledges other factors including: allocation concealment; blinding; attrition rates; imprecision; reporting bias; consistency in results across studies; and the generalisability of evidence. The GRADE system has been adopted by NICE, and is seen as a more considered and sophisticated approach over traditional hierarchies (Nutley, Powell and Davies 2013).

In contrast, there is a substantial body of literature critiquing the simplifications and problematic assumptions in the use of such narrow definitions of evidence. Nutley, Powell and Davies (2013) highlight that the use of such hierarchies becomes more questionable as the intervention being considered becomes more variable, complex, and context dependent. Worrall (2010) has critiqued the RCT approach from a scientific standpoint, highlighting the ineffectiveness of randomisation. Moving beyond EBM to wider health policy, Cookson (2005) notes that policy making is a fundamentally different type of activity to clinical practice, and that there are differences both in the types of evidence it is appropriate to use and its impact.

4.4 The way in which evidence is incorporated into the policy making process

4.4.1 Health policy in England and resource allocation

Within England, health policy is determined by the National Health Service (NHS), which was inaugurated in 1948. This signified the effective nationalisation of health care and meant that the state took over the role of determining health policy, controlling the allocation of resources and ordering priorities (Allsop 1995). Pascall (2007), Webster (1998) and Powell (1997) provide analysis of the progression of the NHS and health policy in England. Suffice to say that the NHS model of health care, financed by central taxation and providing treatment free at the point of use, demanded greater recognition of the classic economic dilemma between potentially unlimited wants and the scarcity of resources (Bryan, Williams and McIver 2007). In 1996 the Department of Health defined the primary purpose of the NHS as “to secure

through the resources available the greatest possible improvement to the physical and mental health of the nation” (Department of Health 1996). Overall, in terms of the level of funding for health care, NHS spending was £100.2bn in 2009/2010⁴⁹ (HMT PESA, p27), accounting for a large proportion of the Government’s £669.26bn total expenditure over this time frame (HMT PESA 2010, p118). Yet, difficult (clinical) resource allocation, rationing, and priority setting questions have been raised for decision making at all levels (Dobrow, Goel and Upshur 2004) against a backdrop of rising demand, changing need, efficiency drives and economic pressures.

One response to these challenges has been a growing interest in EBM and the role of *clinical and cost-effectiveness* as well as the systematic assessment of actual health outcomes. This has given rise to the prominence of the discipline of health economics (Wagstaff and Culyer 2012). Pascall (2007, p442) notes that “managers have an interest in eliminating ineffective treatments in order to make budgets go further and to raise the quality of care.” Of particular significance is the establishment of the National Institute for Health and Care Excellence (NICE),⁵⁰ by government, to produce guidelines on the clinical and cost effectiveness of services (Baggott 1998, p56-7) and which works in partnership with other bodies⁵¹ to “get evidence into practice” (Alliance for Useful Evidence, p9).

4.4.2 The NICE evaluation model

Analyses within the literature have highlighted the importance of independent ‘evidence institutions’ such as NICE in playing a ‘brokering role’ and mediating between the generation and use of knowledge (Hallsworth, Parker and Rutter 2011; Rutter 2012). Jones (2009) notes that such mediation may include processes of: communication; interaction and exchange; intermediaries/credibility; and the demand for ‘knowledge’ amongst policy makers. The Alliance for Useful Evidence report (2014, p9) highlights that NICE creates a “pull” for evidence. Taking this further, Ferlie et al.

⁴⁹ The year 2009/2010 has been used so that a direct comparison can be made with spending on regional policy.

⁵⁰ The National Institute for Clinical Excellence (NICE) was established in 1999. NICE was renamed the National Institute for Health and Care Excellence on 1 April 2013, reflecting the broadening of its remit across health and social care as well as medical care.

⁵¹ Such as the National Institute for Health Research, the Medical Research Council and the university sector.

(2013, p28) contend that the establishment of NICE effectively “*institutionalised*” EBM at the national level. Indeed, a decision in 2003 meant that the NHS was “legally obligated to provide funding for treatments and medical procedures recommended by NICE's Health Technology Assessment (HTA) appraisal board” (NICE 2013b), effectively dictating priorities throughout the UK (Sorenson, Drummond and Kanavos 2008).

HTA essentially embeds the *peer-review* of evidence into NICE decision making processes through the use of expert opinion, public engagement, stakeholder consultations and contestability mechanisms. Engagement with wider social values and interests has been highlighted as a key lesson from NICE processes (Alliance for Useful Evidence 2014). As discussed in detail within the guidance (NICE 2013b), HTA groups are commissioned by NICE including consultees (including manufacturers of the product), commentators (including other manufacturers), clinical specialists, patient experts and commissioning experts. During the assessment phase, an independent academic centre synthesises and analyses all published evidence on the intervention and prepares a report for consideration by the Institute’s Appraisal Committee.

Only a few technologies are selected for HTA appraisal, although in theory potentially any technologies/medications being used in the NHS could be assessed by NICE at some point. Specifically, the stated criteria for HTA selection include: burden of disease; resource impact; clinical and policy importance; presence of inappropriate variation in practice; potential factors affecting the timeliness for the guidance to be produced; and likelihood of guidance having an impact on public health and quality of life, the reduction in health inequalities, or the delivery of quality programs or interventions (NICE 2013b). Therefore, the HTA programme is a relatively reactive process that evaluates new technologies and medications as they emerge on the market and as the evidence base develop.

At first sight, this process may seem to provide an ideal framework for EBPM and for ensuring the validity and relevance of evidence. However, Ferlie et al. (2013, p47) argue that although the NICE process is seen as both clinically and scientifically legitimate with both expert clinical and academic advice as well as patient opinion incorporated; the guidelines embody “bounded pluralism.” Ferlie and McGivern (2014)

suggest the use of this terminology⁵² to explain the power relations between stakeholders, noting that the methodological core resides in a 'bounded elite' represented by advisory groupings of expert clinicians, clinical academics, and health service researchers. Clearly, despite its independence, NICE recommendations are inextricably tied up with political decisions about 'value for money'.⁵³ Harrison (1998) criticises EBM as a solution to resource allocation issues (including those justified by evidence) because it cannot be devoid of political questions such as who benefits. Thus, NICE's work has been controversial and is highly scrutinised by the medical community⁵⁴ and wider society.

4.4.3 Study designs and methods for NICE economic evaluation

NICE has defined a 'reference case' in its methodological guidelines, which "specifies the methods considered by the Institute to be appropriate when preparing submissions for HTA appraisal" (NICE 2013b). Economic evaluation, in particular cost-effectiveness analysis, has been the 'centrepiece' of the NICE resource allocation model (Drummond et al. 2005). In the sections that follow, the methods detailed in the NICE reference case for producing evidence on outcomes, costs, cost-effectiveness, and analysis of confidence in the data, are briefly summarised and expanded upon. The structure of this (brief) analysis is mirrored in the next chapter focused on regional policy to highlight key differences between the sectors.

Analysing outcomes

Drawing knowledge from primary studies is referred to as 'first generation knowledge' by Straus, Tetroe and Graham (2013). Within HTA economic evaluation studies it is expected that "all direct health effects should be included in the analysis" (NICE

⁵² The authors suggest that a model of 'pluralism' (i.e. cooperative bargaining between different coalitions of stakeholders across different issues) is too simplistic to account for elite group coalitions which, they argue, tend to dominate decision-making.

⁵³ For instance, NICE are currently looking at the principles used in relation to the social value of interventions (NICE 2013d). There has been some controversy around this as there is political pressure for social value to be focused on the ability of an individual to be productive in the market. However, this would discriminate against younger people/older people, etc.

⁵⁴ For instance, the British Medical Journal publishes an online editorial commentary series titled 'Controversies in NICE guidance' (www.bmj.com) to highlight failures, success and controversy within NICE guidelines.

2013b). Nutley, Davies and Walter (2002, p3) note that “what counts as a ‘desired outcome’ is readily understood” in medicine. Indeed, clinical effectiveness is usually focused on reductions in mortality, morbidity and clinically important changes in health outcomes for individuals (such as self-reported pain, quality of life and function). Experimental study designs, in particular the use of randomised controlled clinical trials (RCTs), are predominantly used to collect data on such parameters and usually intermediate and final endpoint data are used to measure intervention effectiveness (NICE 2013b). Bryman (2012) describes the basic principle as being that under controlled conditions, there is direct comparison of two or more therapeutic regimens (one of which may be a traditional treatment, a placebo, or the exclusion of active treatment). Subjects are randomly allocated so that extraneous factors that would have affected outcomes for both groups are automatically stripped out and any differences are attributed to the effect of the intervention. This avoids the danger that the results may be biased by other, possibly unobserved, differences between the underlying characteristics of the two groups.

Modelling approaches may also be used to estimate or simulate the effects of a clinical trial when experimental approaches are not feasible (Holford et al. 2000). “When technologies are being compared that have not been evaluated within a single RCT, data from a series of ‘pairwise head-to-head RCTs’ should be presented together with a ‘network meta-analysis’ if appropriate” (NICE 2013b). Glenny et al. (2005) discuss that such modelling of different treatments via ‘indirect comparison’ (i.e. adjusted according to the results of their direct comparison with the common control) enables the strength of the randomised trial to be preserved to a degree. In addition, Reeves et al. (2011) discuss many types of non-randomised *quasi-experimental approaches* that may be used within health including: cohort studies; case-control studies; controlled before-and-after studies; interrupted-time-series studies; and controlled trials that use inappropriate randomisation strategies (sometimes called quasi-randomised studies). Such study designs may explore implementation and operational issues (Reeves et al. 2008). However NICE guidelines (2013b) point to the potential biases in such approaches.

Finally, NICE guidelines (2013b) stipulate that “health effects should be expressed in terms of QALYs” (Quality Adjusted Life Years i.e. a measure of life expectancy and the

quality of the remaining life-years). Phillips (2009) explains that “QALYs are used as a common currency to assess the extent of the benefits gained from a variety of interventions in terms of health-related quality of life [HRQoL] and survival for the patient.” Much research has been conducted on HRQoL and the ‘EQ-5D’ is the preferred measure for adults (NICE 2013b). The EQ-5D comprises five dimensions of health: mobility, ability to self-care, ability to undertake usual activities, pain and discomfort, and anxiety and depression. For each of these dimensions the EQ-5D has three levels of severity (no problems, some problems, severe problems) (EuroQol 2014). QALYs essentially place a weight on time in these different health states.

Systematic reviews

Drawing knowledge from the aggregation of existing knowledge is referred to as 'second generation knowledge' by Straus, Tetroe and Graham (2013). They note such approaches include: systematic reviews; meta-syntheses; scoping reviews; and realist reviews (Straus, Tetroe and Graham 2013). In the literature, such reviews are proposed as a means of bridging the gap between research and decision making, going beyond the raw data (Lavis et al. 2003). For NICE decision making, the analysis of clinical effectiveness is expected to be based on systematic review data from “all relevant studies of the best available quality” (NICE 2013b). RCTs are considered to provide the most “valid evidence of relative efficacy” for systematic review (NICE 2013b). Specific guidelines for systematic reviewing have been developed (Higgins and Green 2011; Centre for Reviews and Dissemination 2008). Within health research a number of initiatives have sought to systematically review the existing evidence base in England such as the ‘Database of Abstracts of Reviews of Effects’ (DARE), the ‘UK Centre for Reviews and Dissemination’ at the University of York, the ‘Cochrane Collaboration’ (health), and ‘Campbell Collaboration’ (social science). However, research synthesis approaches have been criticised. Greenhalgh and Russell (2006) point out that, although judgements are needed to undertake evidence syntheses, there is an assumption that they are ‘technocratic’ and hence can be unbiased in nature through the correct application of the appropriate methodological and evaluative toolkit. They term this the ‘Cochrane inspired myth’. Another key issue is that systematic reviews become out of date quickly as new evidence is produced (Shojania et al. 2007).

Analysing costs

According to NICE guidelines (2013b), “the expected value of each component of cost and expected total costs should be presented” and costs should relate to NHS resources. Market prices (public list prices) of drugs, medical devices etc. should be included, and if there is no competitive market, scales of charges or fees or other forms of administrative reimbursement may be used (e.g. primary care drug tariffs). Healthcare Resource Group (HRG) codes have been developed as standardised reference costs for particular NHS procedures (HSCIC 2014). NICE (2013b) notes that applying HRG costs can “reduce the need for local micro-costing (costing of each individual component of care related to the use of a technology).”

Cost-effectiveness analysis

“Cost-effectiveness (specifically cost-utility) analysis is the preferred form of economic evaluation” for HTA (NICE 2013b). In their seminal paper, Weinstein and Stason (1977) detail the key components for undertaking such analysis. Overall, the aim is “to establish whether differences in expected costs between options can be justified in terms of changes in expected health effects” (NICE 2013b). Modelling methods are used for most technology appraisals, and the guidelines specify the expectation of ‘high quality models’ by the Institute (NICE 2013b). Of interest is that details of the data inputs and any underlying assumptions are expected to be provided for peer-review. In such models, QALYs are combined with the relative cost of treatment to form an ‘Incremental Cost-Effectiveness Ratio’ (ICER) (Folland, Goodman and Stano 2010) which indicates “the ratio of expected additional total cost to expected additional QALYs compared with alternative treatment(s)” (NICE 2013b). It is expected that model estimates will also be reported separately for all relevant subgroups of patients (NICE 2013b).

For HTA, Hounton and Newlands (2012, p1) point out that most new interventions are “likely to be more effective and more costly because breakthroughs in medical procedures and new technologies are typically more expensive than existing practices...there is a need to estimate the maximum society is willing to pay for an additional unit of health gain.” NICE does not set the budget for the NHS, but it has an *allocation threshold* (McCabe, Culyer and Claxton 2008) whereby “the maximum

acceptable ICERs are £20,000-£30,000 per QALY gained” (NICE 2013b). The allocation threshold, effectively constituting explicit national rationing, is controversial and has faced extensive criticism. Culyer et al. (2007) argue that it is not ‘constitutionally appropriate’ for NICE to set such a threshold and that instead the task for NICE should be to build understanding of the most appropriate threshold. The ‘House of Commons Health Select Committee’ stated in 2008 that the “... cost-per-QALY [NICE] use to decide whether a treatment is cost-effective is of serious concern. The threshold it employs is not based on empirical research and is not directly related to the NHS budget” (Great Britain 2008).

Sensitivity analysis

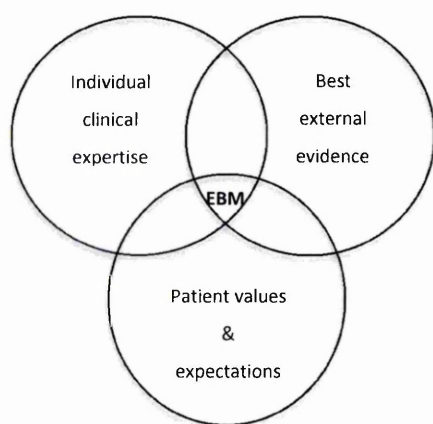
The expectation that uncertainty in the data will be explored is an important component of NICE guidelines (NICE 2013b; Claxton et al. 2005) and sensitivity analysis is used to indicate the robustness of a study. Two types of sensitivity analysis are generally applied and reported upon. ‘One-way Sensitivity Analysis’ can be used to assess the impact that changes in data inputs have on the model’s results (Meltzer 2001). ‘Probabilistic Sensitivity Analysis’ (PSA) can be used to assess multiple sources of uncertainty (Briggs 2000) and to quantify the level of confidence in the model’s results through reporting confidence intervals (Briggs, Claxton and Sculpher 2006). Then ‘Cost Effectiveness Acceptability Curves’ (CEAC) can be used to visually illustrate for decision makers the uncertainty surrounding the estimate of cost-effectiveness in relation to the allocation threshold (Fenwick, Claxton, and Sculpher 2001).

4.5 Other factors besides evidence which affect the way policy is made

A distinct and evolving consideration in EBM has been the role of patient engagement and the ‘human aspects of care’ (Greenhalgh, Howick and Maskrey 2014). Pascall (2007, p419) comments that traditionally “patients had little role in NHS decision making, and were seen as having little role in their own health care.” But over time people accessing health care have developed “greater expectations of choice and control as consumers of services” (Pascall 2007, p420). A particularly contentious issue has been spatial variations, termed ‘postcode lotteries’ in the press, for the prescribing of effective treatments (Pascall 2007).

The role of 'patient values and expectations' form an important component within the EBM model alongside 'best external evidence' and 'individual clinical expertise'. These three elements form the well-known 'EBM triad' as illustrated in Figure 12. The three circles illustrate the distinct but overlapping sources of information that might be used when making clinical decisions (Satterfield et al. 2009). Guyatt et al. (2008, p5) argue that EBM must always consider the patient's values to weigh up "the benefits and risks, inconvenience, and costs associated with alternative management strategies." Taking a broader view, Jones (2009) discusses the concept of 'participatory' (civil society) knowledge which he explains encompasses the voice of the people through civil society organisations, such as through protest groups and lobbying.

Figure 12: The Evidence Based Medicine Triad



Source: adapted from The Cochrane Collaboration 2014

As evidence of patient values and expectations tends to be self-reported and qualitative in nature, there has been a tension. Davison (2013) also argues that the higher value placed on particular forms of information means that when integrating multiple types of evidence, some groups (often those from lower socio-economic backgrounds) are marginalised within the policy making process. Overall, EBM literature and training programs have tended to remain dominated by positivist applications of science with the focus on medical informatics, clinical epidemiology, biostatistics, and critical appraisal (Satterfield et al. 2009).

How precisely different types of evidence, in particular patient values (participatory knowledge) and clinical expertise (expert knowledge) are considered and used by individual decision makers is a question that remains to be addressed in the literature.

Rutter, Hawkins and Parkhurst (2013) reflect that this is particularly interesting given the emphasis and importance that many papers place on such knowledge. Drawing upon work by Pawson et al. (2003) there is a need to weight the role of 'patient values and expectations' (user knowledge) against 'clinical expertise' (organisational and practitioner knowledge) and 'best external evidence' (research and policy community knowledge) in decision making. Similarly, Lin (2003) describes the process of health policy making as one of balancing 'competing rationalities', whereby decision makers must consider arguments of 'technical rationality' (including technical evidence) alongside 'competing political rationality' (what is politically expedient) and 'cultural rationality' (broader social values and understanding). In their paper drawing on the parallels between EBM and EBPM, Dobrow, Goel and Upshur (2004) highlight the importance of the decision making context and individual agency (personal factors) for this.

4.6 The extension of EBM to health and social care

As part of the biggest ever reform of the NHS, the Health and Social Care Act (Great Britain 2012) extended NICE's remit to social care as well as health and clinical care, signifying an increase in the sphere of influence of EBM. However, Greenhalgh, Howick and Maskrey (2014) have recently questioned whether EBM is a 'movement in crisis'. Likewise, Spence (2014, p1) has argued that EBM is 'broken', claiming it leaves "no room for [clinical] discretion or judgment." Satterfield et al. (2009) agree to an extent, noting that the role and value of practitioners and their expertise in EBM is unclear, adding that not enough attention is paid to patient preferences or to contextual factors and resources. In extending the influence of EBM to wider health policy, (i.e. from the micro, clinical level to the macro, policy level), the decision making context becomes more complex (Dobrow, Goel and Upshur 2004). Byford (NICE 2013c, p21) highlights the challenges: there is a difference between populations; between service users; carers; and a wide range and overlap of providers and unpaid carers. In turn, Pawson (NICE 2013c, p20) argues that "as the interventions, programmes, policies and services under review become more complex - so does the challenge of uncovering valid and reliable evidence."

NICE's HTA evidence review process puts great value on a high quality but narrow evidence base, focused on clinical expertise, RCTs, and systematic reviews.⁵⁵ Ho, Peterson and Masoudi (2008) counter that evidence hierarchies are inappropriate for many health outcomes research questions, and that RCTs are often not feasible or appropriate for complex, frontline work within health.⁵⁶ Thus there will be a narrow evidence base for NICE to draw upon when developing its guidelines (Knaapen 2013; NICE 2013b) and health policy decisions are based on more than just evaluations of effectiveness (Petticrew and Roberts 2003; Abeyasinghe and Parkhurst 2013).

It has been stressed that RCTs and systematic reviews are costly to carry out and that, as the pharmaceutical industry funds most RCT research (see Buchkowsky and Jewesson 2004), this leads to "commissioning bias" (Spence 2014, p1). Every-Palmer and Howick (2014) have therefore argued that EBM is failing due to biased trials and selective publication. Barnes and Parkhurst (2014) explain that policy decisions may become biased towards those issues that are amenable to RCT design (i.e. treatment effectiveness), and away from complex social issues harder to evaluate by such methods (i.e. efforts to address the social determinants of health). Overall, Byford (NICE 2013c, p21) has argued that, for the NICE evaluation model to be applied more widely, a more flexible approach to economic assessment is needed alongside communicating methodological limitations and addressing gaps in the evidence base.⁵⁷

NICE has acknowledged many of these issues (Alliance for Useful Evidence 2014). In terms of changing practice, the NICE paper 'Social value judgements' (2013d) sets out the Institute's approach when there is a paucity of high quality evidence, and this guidance is currently being revised (to be completed in early 2015). In addition, NICE does have considerable experience in public health where similar issues exist with regards to the limitations of the evidence base. However, overall there has been a call for a movement towards the more *appropriate* use of evidence for policy decisions,

⁵⁵ NICE's broader guidance programmes (beyond HTA) aim to utilise the 'best available evidence' (which in many cases is expert opinion rather than RCT's) (NICE 2013a).

⁵⁶ Indeed, the challenges and potential for NICE to work in the area of social care have been highlighted before (Gould and Kendall 2007).

⁵⁷ Part of NICE's work is to review knowledge gaps (for instance, NICE has contributed towards setting up a 'Database of Cancer Uncertainties' (DoCU).

rather than simply calling for the increased uptake of particular forms of evidence (Abeyasinghe and Parkhurst 2013).

4.7 Conclusion

This chapter has contributed in a number of ways to address the first research question. It was found that within health policy EBM has grown to become a large and powerful movement and its sphere of influence has expanded far beyond its origins in internal medicine. Therefore there has been critical scrutiny and debate in the literature over the different types and credibility of evidence and the balance between individual policy maker expertise (organisational/practitioner knowledge), patient values and expectations (user knowledge), and external evidence (research/ policy community knowledge). This chapter has demonstrated that the establishment of NICE has provided an institutional process, political backing and a legislative framework to incorporate evidence into policy making. Although open to strong criticism, this has enabled a process of stakeholder engagement, peer-review and the setting of clear guidelines to evidence producers. The Institute has also created demand for economic evaluation evidence and the use of 'modelling', research synthesis and sensitivity analysis.

However, this analysis identified that the EBM rational choice model fails to engage with the political nature of decision making; meaning that the actual delivery of health care is likely to show some differences in practice. This was found to be particularly the case when moving beyond the micro to the macro decision making level, where context becomes increasingly important. It can be deduced that the application of a "one size fits all" (Goodman 1999, p250) NICE (HTA type) evaluation model, which fails to take into account wider social values and interests, is therefore less reconcilable within the complex decision making environments for wider health and social care policy. Accordingly, consumers of evidence need to be cautious of the false sense of certainty which could be created by the NICE process and understand the limitations of the evidence base to prevent misled policy formulation (Bovaird 2014).

The structure of this chapter focused on health policy and EBM provides the foundation for the next one, focused on regional policy and impact evaluation.

Chapter 5

Regional Policy & Impact Evaluation

5.1 Introduction

This chapter builds on Chapter 4, drawing upon methodological guidelines for economic evaluation and the academic and policy literature to contribute towards addressing the first research question. This chapter provides the foundation for the analysis conducted in Chapter 6 whereby EBM approaches are used as a yardstick against which wider social policy is assessed and parallels are drawn between the practices of EBM and EBRPM. The overarching Evidence Based Policy Making (EBPM) debates will once again be explored, but focusing on the case study of regional policy and impact evaluation. As such the following will be discussed: the hybrid of ‘top down’ and ‘bottom up’ evidence; the practice of regional policy evaluation in England by the institutions charged with its implementation; the role of other factors besides evidence such as the inherently political nature of regional policy; and the influence of decision support. To finish, the challenges of applying an EBPM approach to regional policy making will be explored by focusing on the Regional Development Agency (RDA) national impact evaluation.

5.1.1 Theoretical background

This chapter predominantly focuses on debates around evidence types and the role of research credibility. Therefore the theoretical background for this chapter is provided in Chapter 2 (section 2.3).

5.1.2 Scope: Regional policy in England

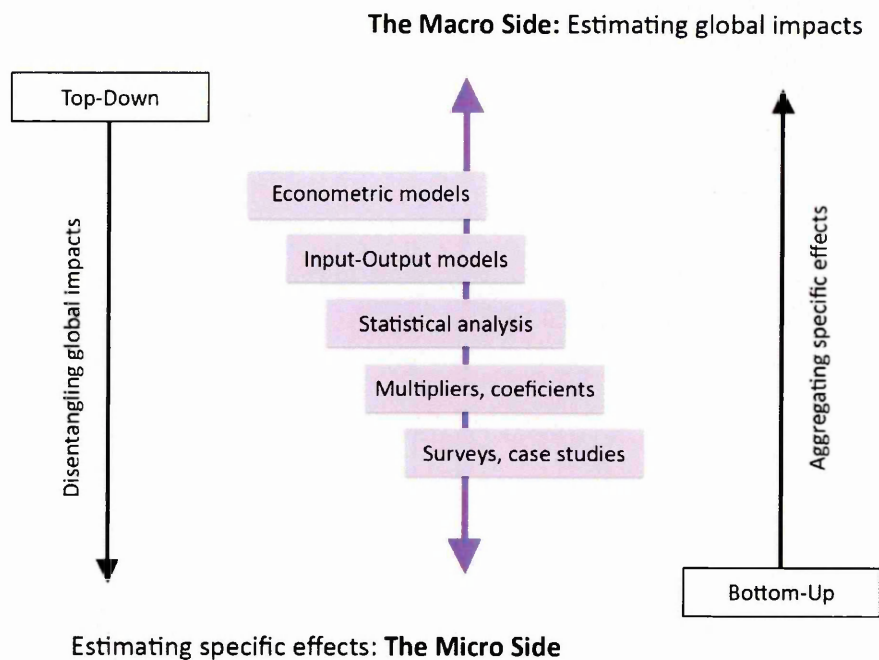
The background to regional policy, including the rationale, history, aims, institutions and participants was considered in detail in Chapter 1 (section 1.2 and 1.3) and the current situation for regional policy evaluation discussed (section 1.4). This chapter focuses on the development of regional policy evaluation and EBRPM. It should be borne in mind that regional policy has subsets and related fields (such as small business policy, skills policy, and infrastructure planning), and a number of subnational institutions deliver economic policies. Armstrong and Wells (2006) highlight that

regional policy has spread (mission creep) into urban and community policy arenas. However the scope of this research focuses on regional policy as delivered through the RDAs. In particular, the RDA evaluation model is focused upon and the challenges of extending an EBPM approach to regional policy are then briefly examined. The purpose of this approach is to mirror the analysis undertaken in Chapter 4 and to highlight the key differences across the regional policy and health policy sectors. The discussion that follows is relatively high-level and sets the background for the comparative literature review undertaken in the following chapter (Chapter 6).

5.2 Evidence types and the role of research credibility

There has not been a formal 'evidence based movement' within regional policy evaluation practice equivalent to EBM within health policy. Despite this, Wells (2007, p27) identifies that evaluation evidence "has become a more widely accepted part of the policy making process, more frequently and knowledgably used by central government and local and regional agencies." In terms of appraising the quality of project/programme evaluation, explicit hierarchies of evidence quality have not been commonly applied in practice or critically examined in the regional policy literature. Having said that, it could be argued that a classification of evaluation study designs (Figure 13) that was presented in the EU MEANS guidance (European Commission 1995), indicating 'top-down' though to 'bottom-up' approaches, suggested an *implicit* evidence hierarchy.

Figure 13: MEANS classification of the types of evaluation (1995)



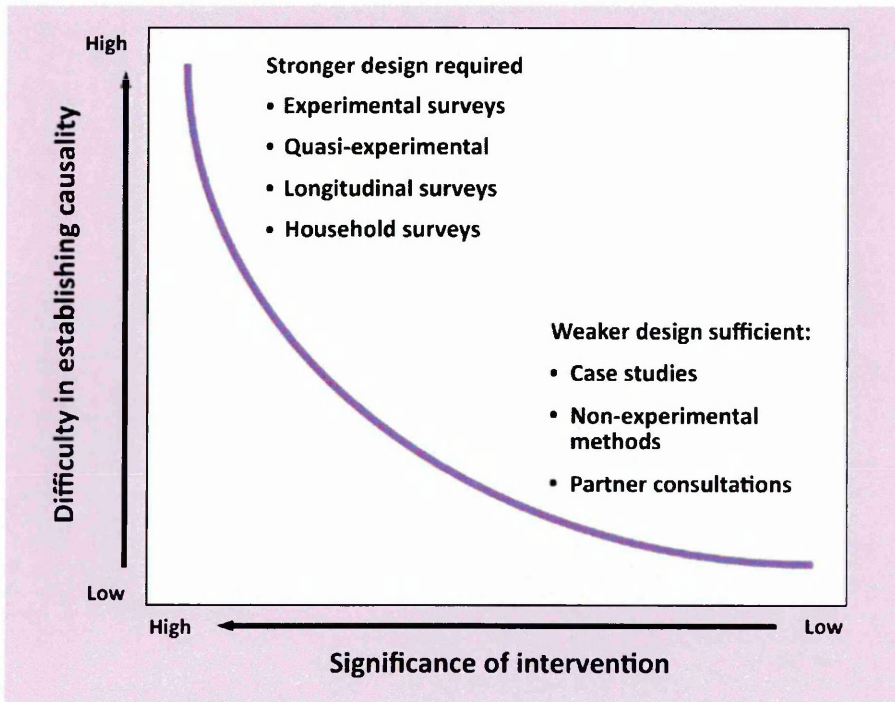
Source: European Commission 1995, p24

Top-down methods draw upon secondary data sets (i.e. regional unemployment time series, or industrial location cross-sectional data) to “estimate impacts on indicators such as employment and value added” (Armstrong and Wells (2006a, p857). In contrast, bottom-up approaches “draw upon primary data collected directly from the beneficiaries of regional policy (i.e. individuals and enterprises) using survey methods and case studies” (Armstrong and Wells (2006a, p857). (Quasi) experimental evaluation methods (e.g. regression or matched-pairs analysis) are in the middle of this spectrum (Isserman and Rephann 1995). Collectively these approaches are termed ‘method-based evaluation’ (MBE).

The MEANS handbook (European Commission 1995, p24) stated that the five methodologies presented in Figure 13 covered the full range between “classical” micro-approaches and the “sophisticated end of the macroeconomic scale.” The term ‘sophisticated’ could have been interpreted at the time as symbolising more progressive approaches than more traditional ‘classical’ approaches. In turn an implicit evidence hierarchy could have been inferred whereby macro, quantitative approaches (such as econometric modelling) are placed at the top, and micro, more qualitative

approaches (such as case studies) are placed at the bottom. Although the MEANS guidance was initially published in 1995, more recent RDA evaluation guidance (DTI 2006, p46) presented a similar implicit evidence hierarchy (Figure 14). In this classification, evidence quality was more directly inferred, with “stronger” study designs (experimental surveys and quasi-experimental approaches) compared to “weaker” study designs (non-experimental methods and partner consultations).

Figure 14: IEF classification of the types of evaluation (2006)



Source: DTI 2006, p46

In terms of the philosophical foundations underpinning such regional evaluation studies, two papers by Armstrong and Wells (2006a; 2006b) are of interest. They describe the evolution of regional policy evaluation (particularly Structural Fund evaluation) in the UK. They explain that early regional policy evaluation was predominantly positivist in nature, rooted in orthodox economic theory with a focus on statistical techniques. Indeed, Garretsen et al. (2013, p182) highlight the ground-breaking work of Moore and Rhodes (1973) as a ‘turning point’ for regional policy evaluation in identifying a credible and robust “counterfactual with which to establish the impact of policy on new investment and job creation in the British Assisted

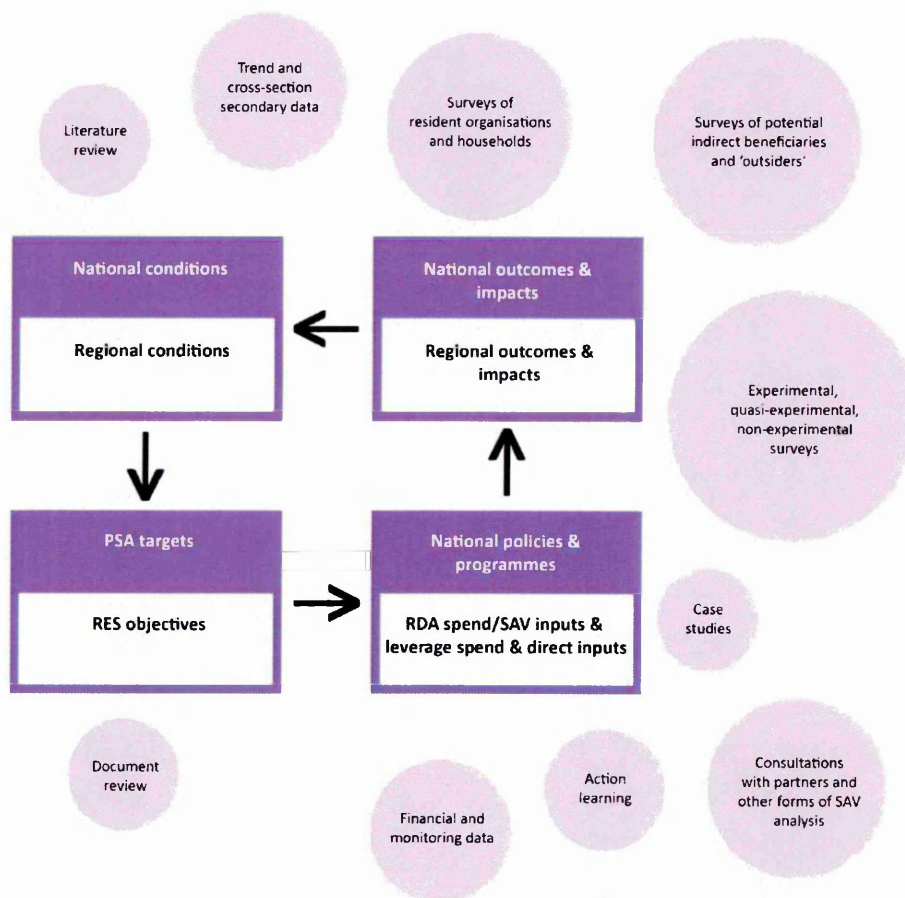
Regions.” Armstrong and Wells (2006b, p264), explain that a positivist approach was then “carried over into the Structural Funds evaluations” after 1989.

The academic literature focused on top-down methods became increasingly sophisticated over time, with research drawing upon time-series regression (Wren and Taylor 1999), input-output analysis, and computable general equilibrium models (Gillespie et al. 2001). In addition, over the last three decades, the use of quasi-experimental study designs increased substantially in the regional science literature (Feser 2013). In particular, Isserman and colleagues have been identified for their notable work on comparison group designs, and experimentation in regional research settings (Feser 2013; Markusen 2015). However, Armstrong and Wells (2006a) note that in practice, bottom-up techniques began to form the ‘backbone’ of the Structural Fund evaluations from the late 1980s onwards. They explain that such bottom-up approaches were still positivist in nature, in that attempts were made to aggregate micro-level data and to identify a counterfactual (i.e. by directly asking questions to those surveyed to elicit ‘deadweight’ and ‘displacement’ effects). However, it could be contended that the focus on bottom-up techniques has meant that there has not been the demand to translate the increasingly sophisticated econometric and statistical techniques presented in the academic literature into practice, as such techniques have tended to be more closely aligned with top-down methods.

In addition, Armstrong and Wells (2006a) note that, although MBE remains a major part of the Structural Fund evaluations, its positivist foundations have been challenged by constructivism and realistic approaches. In particular, realistic approaches have become more widely adopted and Garretsen et al. (2013, p182) note that “by the late 1990s New Labour emphasised that spatial interventions should be underpinned by a robust theory of change, relevant logic chains and a focus on outcomes.” Given that many contemporary regional policies focused on social inclusion and community empowerment, more participatory approaches to evaluation practice were emphasised (Diez 2002). Focusing on the RDA evaluation model, a diagram in the Impact Evaluation Framework (DTI 2006) highlighted the potential number of co-producers of RDA evaluation evidence (Figure 15). The following actors are identified: national government; RDAs; beneficiaries; non-beneficiaries; resident organisations and households; indirect beneficiaries and ‘outsiders;’ and partners.

In Figure 15, a *hybrid* RDA evaluation model is indicated, integrating various forms of evidence and processes of collating evidence from independent expert through to action learning approaches (DTI 2006, p42-45). This may have reflected New Labour's overarching approach to EBPM at the time which was focused towards policy learning (see Wells 2007). Of course this evaluation model raised the challenge of assessing the quality, relevance and comparability of the evidence produced; particularly when RDA evaluation production had not been based upon a process of peer-review akin to NICE Health Technology Assessment (HTA). Wells (2014, pers. comm.) reflects that ultimately "the RDA evaluation model tried to keep many actors happy, which was a high transaction cost business."

Figure 15: RDA evaluation framework and methods of evaluative data collection



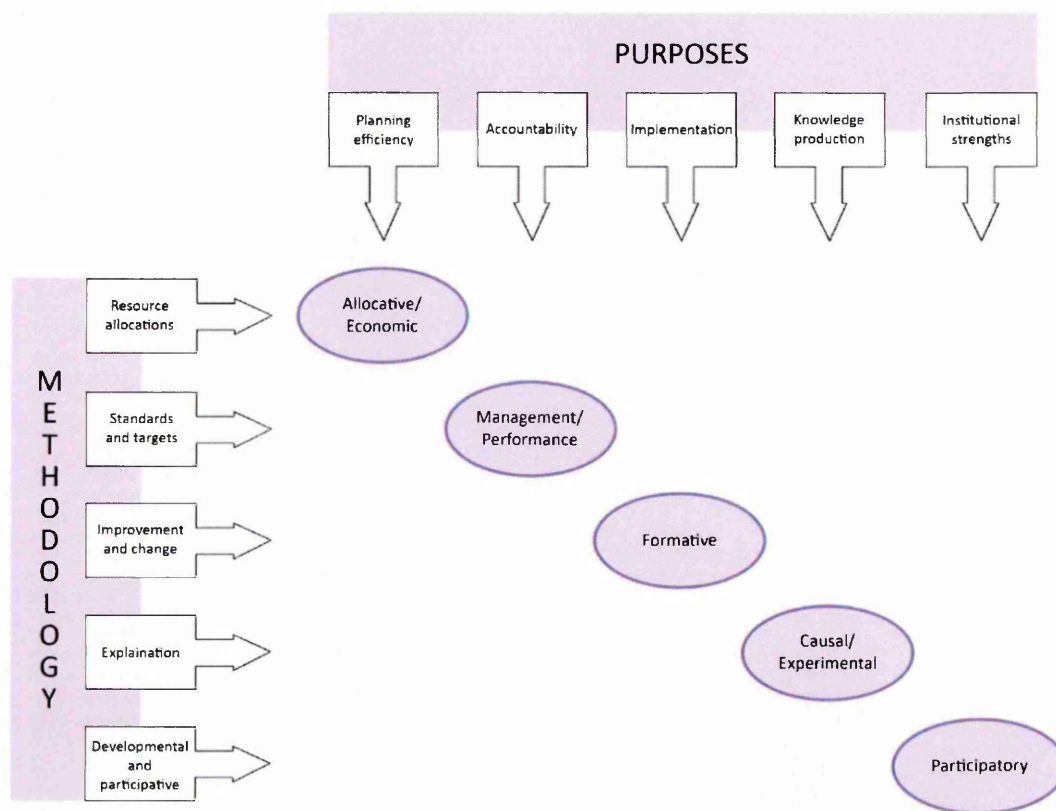
Source: DTI 2006, p41

Returning to the concept of implicit evidence hierarchies, the Tavistock Guide (Tavistock Institute 2003) attempted to identify a more nuanced classification of

evaluation evidence than those identified in Figures 13 and 14, defining different methodological positions in relation to the different purposes of evaluation (Figure 15).

Reflecting upon Figure 16, Armstrong and Wells (2006b, p267) identify a division of approaches between programme and project level evaluations. They note that “the big set-piece reports” for the Structural Fund evaluations, inferred to be those mostly undertaken by central EU economic agencies, tended to use the first three types of evaluation shown (i.e. allocative/economic; management/performance; formative), whereas individual project level evaluation, inferred to be those mostly undertaken by local and regional agencies, tended to use the final two types of evaluation (i.e. causal/experimental; participatory).

Figure 16: The Tavistock Guide classification of the types of evaluation



Source: The Tavistock Guide (2003) cited in Armstrong and Wells (2006b, p267)

In practice the credibility of regional policy project/programme evaluation outputs has faced criticism (Rhodes, Tyler and Brennan 2005; O'Reilly 2007; Tyler and Brennan 2007) with arguments pointing towards weaknesses in research design as well as a lack

of evidence on the impact of regional policy interventions on key outcomes (Rhodes, Tyler and Brennan 2005; O'Reilly 2007). Armstrong et al. (2002) suggested that a gap had emerged between the practice of evaluation and the needs of regional policy making and practice.

5.3 The way in which evidence is incorporated into the policy making process

5.3.1 The RDA and Central Government evaluation model

The RDAs were financed by national government public funds⁵⁸ via the creation of a 'single pot' of RDA funding.⁵⁹ There was a high degree of flexibility, enabling funding to be targeted towards the investments the RDAs prioritised for each region (YF 2009). Overall, in terms of the funding available, RDAs' combined single pot budgets were £2.26bn in 2009/2010 (GREAT BRITAIN 2015, p18)⁶⁰, accounting for a small proportion of the Government's £669.26bn total expenditure over this time frame (HMT PESA 2010, p118).

There has not been a formal legislative or statutory regulatory framework influencing the supply and use of evaluation evidence within regional policy evaluation practice at a national level equivalent to EBM and NICE within health policy. Although regional policy evaluation in England is long-standing and pre-dates EU accession, the influence of EU regional policy and the subsequent demands to evaluate Structural Fund expenditure from 1989 onwards means that EU economic agencies have been at the forefront of developing methodology for evaluation practice and a common set of guidelines for the monitoring and evaluation of EU programmes within and across regions (Bachtler and Michie 1995). Initially, there was a paucity of guidance setting out the approach for regional policy evaluation in England.

⁵⁸ In addition to European Regional Development Funding (see Chapter 1).

⁵⁹ 'Single Pot' pooled money from all the contributing government departments in the UK: The Department for Business, Innovation & Skills (BIS); The Department for Communities and Local Government (CLG); The Department for Innovation, Universities and Skills (DIUS); The Department for Environment, Food and Rural Affairs (Defra); The Department for Culture, Media & Sport (DCMS); and UK Trade & Investment (UKTI). BIS was the sponsoring department (YF 2009).

⁶⁰ The year 2009/2010 has been used so that a direct comparison can be made with spending on health policy.

Over time the approach was formalised across a number of government publications including: the EC MEANS guidance (European Commission 1995) and post-MEANS guidance on the Evaluation of Socio-Economic Development (EVALSED);⁶¹ the HM Treasury's Green Book (2003b); and the Cabinet Office's Magenta Book (2011). For regional policy evaluation specifically there was the English Partnerships' Additionality Guidance (2008) and also the 3Rs Guidance (ODPM 2004). However, when the RDAs were established, they were given very little guidance as to what was expected by central government when it came to evaluation of projects and programmes (see for example, stage 6 of the Single Programme Appraisal Guidance, DTI 2003). In their later years, the RDAs developed with BIS the Impact Evaluation Framework (IEF) (DTI 2006) and Practical Guidance on Implementing the Impact Evaluation Framework (known as the 'IEF +') (BIS 2009a; 2009b). The guidance attempted to underline the purpose and merits of conducting evaluation in principle and to set the bar as to what was expected of evaluation. In the 'IEF+', a guideline of externally evaluating a minimum of 60 per cent of RDA project/programme spend was stipulated and RDAs were required to submit evaluation plans to BIS each year for peer-review (BIS 2009a).

Until recently,⁶² no independent 'evidence institution' (such as NICE) existed to play a brokering role and to mediate between the generation and use of knowledge for regional policy making. What did occur was a general strengthening of regional policy evaluation co-ordination and resourcing over time (Bachtler 2011, p94). BIS and the RDAs used approaches such as allocating ring-fenced funding for policy evaluation and building into the requirements of funding to delivery partners that policy must be evidence based. Evaluation units were established in the EU Commission and in BIS,⁶³ and all RDAs had established evaluation units/teams between 2007 and 2009 working

⁶¹ As discussed in the EVALSED guidance (European Commission 2013, p1), "Evalsed had its origins in the MEANS programme (Means for Evaluating Actions of a Structural Nature) which started in 1995 and culminated in a 6 volume survey of evaluation in 1999 (no longer in print). In 2004, MEANS was transformed and developed into a website -Evalsed (Evaluation of Socio-Economic Development). Evalsed was further developed in 2007, 2008, 2009, 2012 and, most recently, in 2013."

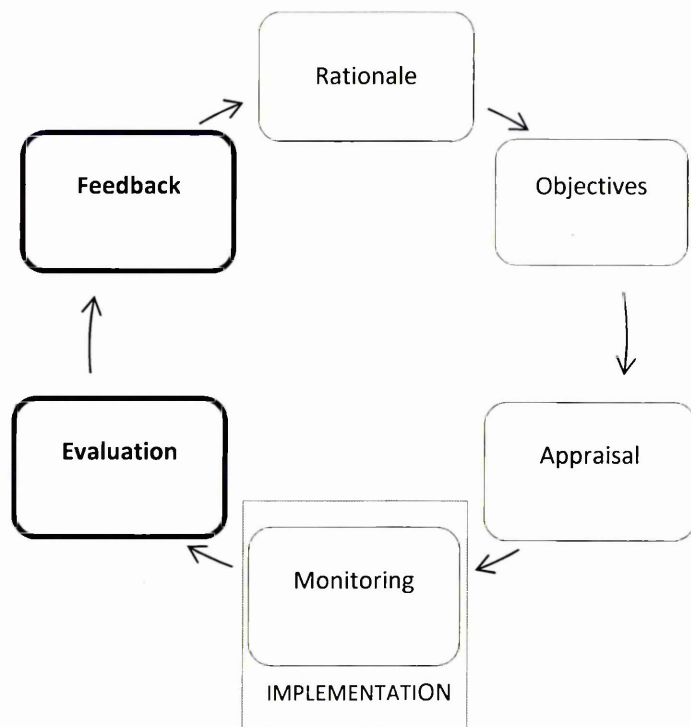
⁶² This refers to the establishment of the 'What Works Centre for Local Economic Growth' in 2013.

⁶³ Strategy Unit's Performance and Evaluation Team within BIS (Bachtler 2011, p104)

alongside an RDA National Secretariat⁶⁴ to commission and produce evaluation evidence.⁶⁵

Despite this focus on the supply side, expectations for how evidence was to be incorporated into policy making processes, and the division between central and regional agencies, have been somewhat unclear. The HM Treasury's Green Book (2003b) recommended a (strategic) framework for the appraisal and evaluation of all policies, programmes and projects (Figure 17). After evaluation there is an expectation that evaluation findings would "feedback" into policy making processes. However operational guidance on how such learning was to be diffused or cascaded down within regional policy making was not formalised centrally.

Figure 17: ROAMEF cycle



Source: The Green Book (HM Treasury 2003b, p 3)

⁶⁴ Interviewee (2014. pers. comm.)

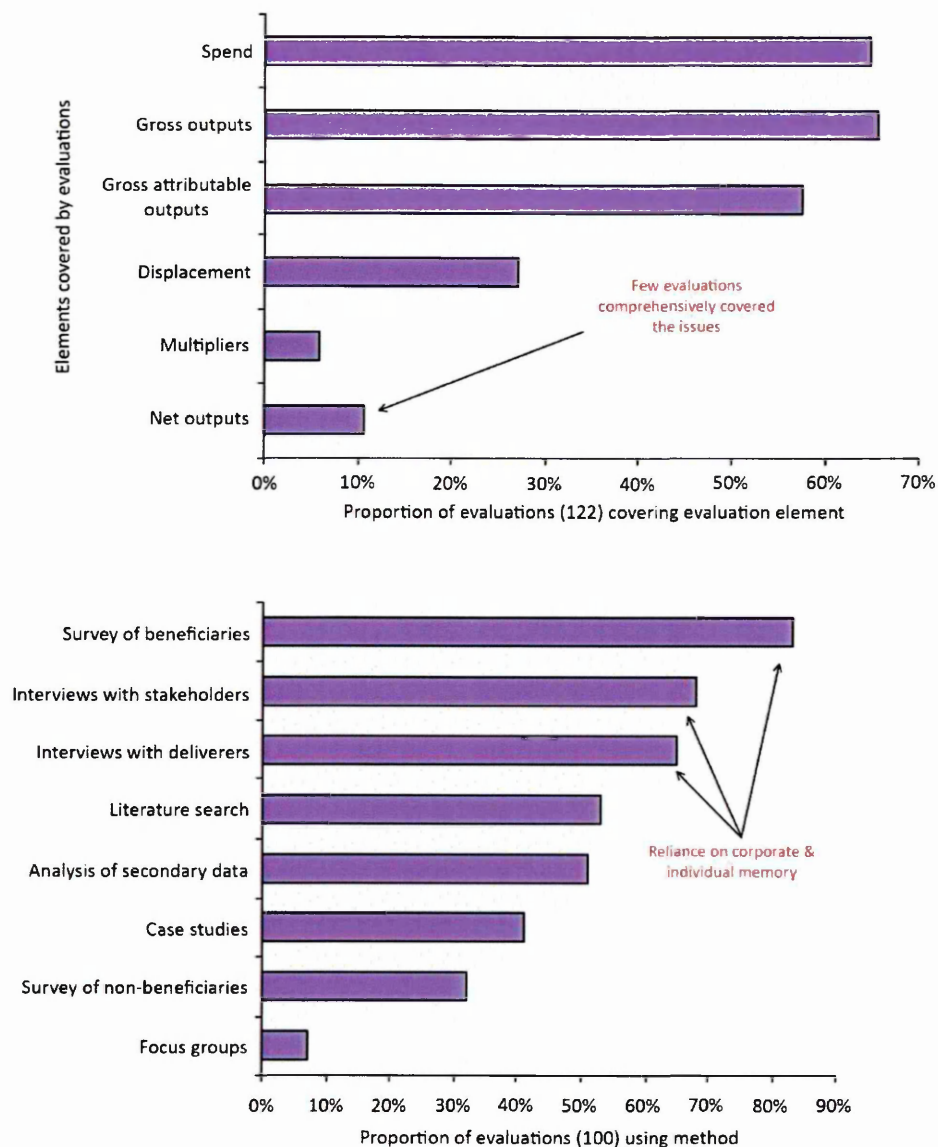
⁶⁵ There was a large variation across the RDAs. Some RDAs established evaluation teams/units in 2004. Sometimes evaluation was headed by Director level staff; sometimes it was headed by middle managers. In some RDAs the evaluation officers acted in a predominantly advisory role and were not necessarily involved in all projects, whereas in other RDAs evaluation was much more centralised.

As highlighted by Figure 15, the multi-level governance structure of the RDA policy/evaluation model was much more complex than the process of Knowledge Transfer and Exchange (KTE) suggested by the ROAMEF cycle (Figure 17). It could be contended that there was not an identified point in the policy process when the integration of the multiple forms of evaluation (rather than monitoring) evidence would be expected to be reported back to central government, certainly not comparable to the NICE or EU evaluation programme model.⁶⁶

A stimulus for change was brought about by analyses in two reports produced by SQW Consulting (cited in Cook et al. 2008), which underpin Figure 18. This work highlighted the weaknesses of the evidence base in demonstrating a rationale for regional policy intervention. In particular, the limited extent to which past evaluations had taken account of the different factors in the assessment of the 'additionality' of an intervention (to identify a counterfactual) and the limited use that had been made of different evaluation methods was evidenced in the reports.

⁶⁶ Ex-ante, mid-term and ex-post evaluation and more recently ongoing evaluations have been stipulated for EU Programme evaluations (European Commission 2006a; 2007).

Figure 18: SQW analysis of evaluation methods and coverage



Source: SQW Consulting, Reviews for UKTI (2005) and the Enterprise Directorate (2006) cited in Cook et al. (2008).

Although the analysis and reporting of the data presented in Figure 18 is open to debate, the implications of the reports are noted by Chadwick, Tyler and Warnock (2013, p844):

It was found that the available evaluation evidence was far too limited; too many evaluations focused on the process issues, were too qualitative in nature and were unable to draw any firm conclusions on impact, often due to inadequate beneficiary data.

In addition, Wells (2007) notes that there was a shift in New Labour's approach towards EBPM more generally, away from policy learning towards policy delivery, with

greater attention placed on 'hard' quantitative and economic analysis. Against this backdrop, in 2006 the IEF (DTI 2006) was published, which sought to establish a consistent framework across the RDA network and which placed emphasis on quantitatively assessing the net economic impacts of interventions. In addition the concept of 'Strategic Added Value'⁶⁷ was added to the RDA monitoring framework to encapsulate the role of RDAs in delivering unquantifiable benefits such as regional leadership and partnership working (YF 2011, p15).

In 2007, the Government commissioned the national RDA impact evaluation exercise from PricewaterhouseCoopers (PwC). The planned approach was to aggregate evaluations deemed to be 'IEF compliant' by PwC and targets were set for RDA expenditure to be covered by compliant, independent impact evaluations (YF 2011, p15). Between 2007 and 2009 over 274 evaluations were commissioned with the results aggregated for the final report (PWC 2009a; 2009b). Reflecting upon this Chadwick, Tyler and Warnock (2013, p845) note that "this assessment did at least provide the application of a common method, which was a major step forward."⁶⁸ The national RDA impact evaluation was the first attempt to provide an independent assessment of the net impact and economic value of RDA interventions (Chadwick, Tyler and Warnock 2013).

5.3.2 Study designs and methods for RDA economic evaluation

Traditionally regional policy appraisal and evaluation tended to focus on the tangible economic benefit of the 'Exchequer cost per job' (Swales 1997). However, the requirement by central government for the RDAs to deliver outputs/outcomes which positively affect GVA growth via the PSA7 target (Chapter 1) shifted the focus somewhat towards measuring (potential) GVA impact (ONS 2010). In addition, the pursuit of economic evaluation, in particular Cost Benefit Analysis (CBA), is demonstrated in the HM Treasury's Green Book (2003b) and is also apparent in European regional policy guidance (European Commission 2006b). However, the IEF

⁶⁷ "Strategic added value was a concept that tried to encapsulate the role of RDAs in delivering unquantifiable benefits such as regional leadership and partnership working" (YF 2011c, p15) i.e. it was the role of influencing others to take action to meet regional and national objectives.

⁶⁸ Lessons learned from the process were then reflected in the 'IEF+' (BIS 2009a; 2009b).

guidance (DTI 2006; BIS 2009a) acknowledged the difficulty of fully monetising all costs and benefits and, rather than (full) economic evaluation, stipulated the need to account for 'net' economic impact via the use of *impact evaluation* to measure net GVA.

There is a longstanding debate as to what should be the key outcome of regional policy. The earliest debates focused on jobs when regions were blighted by high unemployment. Whilst unemployment and indeed hidden unemployment (see for example Beatty, Fothergill and Gore 2012) remain high in weaker regions, in the UK at least the rise of a more flexible labour market has given rise to a concern with underemployment, job insecurity and various symptoms of a low skills equilibrium. The approach of national governments and indeed the European Commission has also shifted. For instance, in the 1990s the focus of EU SFs was far more on job creation, but more recently it has shifted to increasing the level of regional output or regional income. This is a clear corollary of national concerns with raising GDP. However, producing valid GVA data at a regional level in countries such as the UK and in particular England is far from straightforward as Gripaios and Bishop (2006) acknowledge.

Nonetheless GVA measures have come to the fore as a key policy outcome of both domestic and EU regional policy evaluations. However, the limitations need to be acknowledged. The composition of the GVA measure in bottom-up regional policy evaluations is simply put as net additional profits plus net additional wages. Again whilst a useful proxy of national accounts and income (as well as attempting to reflect the value of jobs created for the economy), it appears to neglect what should perhaps be the overriding concern of welfare oriented policies, namely the overall wellbeing of the population. It is easy within GVA analysis to neglect the balance between businesses and employees and the distributional effects of policy. In practice there has remained a strong focus on analysing employment creation and cost per job measures (Bachtler 2011, p104), which may reflect the above concerns surrounding the use of GVA. More recent work by Tyler et al. (2013) has questioned the rise of other measures of regional performance, and thus their use as outcome measures. This includes work on both competitiveness (an early focus of RDAs) and resilience (a theme of LEPs). Others have argued for the consideration of quality of life or wellbeing

measures to capture the hedonic and eudemonic wellbeing of local populations. It is the latter arena where there is some overlap between regional economic policy and health.

In the sections that follow, the methods detailed in the guidance for producing evidence on outcomes, costs, cost-effectiveness, and analysis of confidence in the data, are briefly summarised and expanded upon. The structure of this (brief) analysis mirrors Chapter 4 to highlight key differences between the health policy and regional policy sectors.

Analysing outcomes

As discussed above, the analysis of outcomes has been an evolving consideration within regional policy evaluation methodological guidance. A key issue has been the differentiation of outputs from outcomes both conceptually and practically. The Treasury has attempted to illustrate their differences by giving some examples (see Table 10), which highlight the fact that outputs are specific and outcomes are more vaguely defined improvement.

Table 10: Examples of Outputs and Outcomes from the Treasury

Policy area	Outputs	Outcomes
Job search/ Job Matching	Number of job seekers	Value of extra output, or improvement in efficiency of job search
Development of skills	Number of training places and /or numbers completing training	Value of extra human capital, and/ or earnings capacity
Social outputs; Schools; Health Centres	Exam results (schools), People treated (health centres)	Improvements in human capital (schools); Measures of health gain (health centres)
Environmental improvement	Hectares of derelict land freed of pollution	Improvement to the productivity of the land

Source: HM Treasury (2003, p14)

Rhodes, Tyler and Brennan (2005, p1942) note that evaluations “...tended to be dominated by either discussions of process or a seemingly endless fascination with the outputs produced by policies.” However, Chadwick, Tyler and Warnock (2013, p848) identify that over time there was a re-orientation of emphasis within central government “to identify the key outcomes and impacts of policy, rather than simply to

identify the outputs that regeneration expenditure produced.” The IEF (DTI 2006, p51) signified the implementation of a common approach and the guidance proposed the use of more robust quantitative evaluation methods and forms of analysis (especially for major programmes and projects) including: the use of longitudinal surveys of beneficiaries; surveys of non-beneficiaries; data linking to the ONS; multi-variate analysis of secondary data sources; the pooling of evaluation data and evidence; and the use of intermediate outcome measures. The framework also advocated the need to capture more quantitative and qualitative data, accounting for other significant economic, social and environmental impacts, to support a more sophisticated assessment of impact (DTI 2006, p62-69) and Strategic Added Value (DTI 2006, p19-21). However, it could be argued that the principles and issues to consider were acknowledged in the guidance rather than a prescriptive methodology detailed. For instance, a section of the guidance on how to calculate social and environmental impacts remained uncompleted.

The use of ‘bottom-up’ approaches focusing on primary research, and particularly the use of beneficiary surveys, remained prevalent after the publishing of the IEF. The revised IEF+ (BIS 2009a, p21), strongly supported this stance with beneficiary surveys advocated as a fit-for-purpose method to disentangle the impact of regional policy from other influences (i.e. to measure the ‘counterfactual’). The guidance (BIS 2009b) went so far as to detail standard questions to be used to support the elicitation of ‘deadweight’ and ‘displacement’ effects.⁶⁹ The guidance (BIS 2009a, p21) stated that quasi and experimental approaches were generally not appropriate due to the difficulties in identifying appropriate control groups and due to the cost. If primary research could not be undertaken, the IEF+ recommended the use of secondary data such as information from company accounts, from similar evaluations, or using benchmark factors (see BIS 2009b).

In addition, for the IEF and the RDA national impact report, an exercise was undertaken across the RDA network to categorise programmes and projects into three overarching IEF ‘themes’. These were: business development and competitiveness; regeneration through physical infrastructure; and people and skills. Each of these

⁶⁹ See the additional calculation and the definitions of the distinct elements of the calculation (including deadweight and displacement effects) on page 133.

themes then contained a series of sub-themes (see Appendix 1). Highlighting the disparity across the RDA network, Chadwick, Tyler and Warnock (2013, p845) note that the complexity of categorising RDA activity and expenditure consistently under these three headings was “significant.”

Research synthesis

Research synthesis has been an under-developed approach within regional policy in comparison to health policy. It is noted in the Green Book (HM Treasury 2003b, p 47) that “efforts should be made to disseminate the [evaluation] results widely, and, for this purpose, it may be helpful to use summaries of the main points, and reports which synthesise the results from a number of evaluations with common features.”

Systematic review and meta-analysis are not widely applied or even discussed in the Green Book or IEF guidance. Pawson et al. (2004, p iv) suggest ‘realist synthesis’ as an “approach to reviewing research evidence on complex social interventions to provide an explanatory analysis of how and why they work (or don’t work) in particular contexts/settings.” Yet the application of realist synthesis to regional policy evaluation in practice has been limited. Instead, narrative review (i.e. evidence reviews and evidence assessments) has been the approach undertaken within RDAs historically, whereby there is an intuitive (potentially subjective) aggregation of individual research findings. Wells (2007) noted that there was increasing emphasis placed on such reviews. Indeed, the work of the Office of Project and Programme Advice and Training (OffPAT) is of interest here. The OffPAT e-library provided a shared repository where RDAs could place completed evaluation reports and OffPAT would then produce a short executive summary of each evaluation (see Chapter 7). However, it could be argued that better use could have been made of these evaluation reports to provide syntheses of practical utility.

Analysing costs

The Green Book (HM Treasury 2003b, p101) states that “as many of the costs and benefits should be quantified in monetary terms as feasible.” The guidance states that costs are to be based on market prices and that “wider social and environmental costs and benefits (for which there is no market price) also need to be brought into any assessment” (HM Treasury 2003b, p19). Applying shadow-pricing unit values to

outcome change data has been explored (Wilson 2012). However, the IEF and IEF+ acknowledged that the full quantification of costs (and benefits) is impracticable for regional policy evaluation as the outcomes of regional programmes tend not to have market value.

Impact evaluation

The IEF required a common definition and approach to estimating 'additionality'. Combining advice given in the Green Book,⁷⁰ SPAG⁷¹ and the 3Rs,⁷² the English Partnerships Additionality Guide (English Partnerships 2008) advocated calculating net impact through a standardised approach:

$$\text{Additionality} = \text{Gross Impact} - \text{Deadweight} - \text{Substitution} - \text{Displacement} - \text{Leakage} - \text{Crowding Out} + \text{Multiplier Effects}$$

Source: BIS 2009b, p24

The distinct elements of the calculation are described in this extended quotation by McVittie and Swales (2007, p13):

Gross impact is simply the activity directly associated with the aided project. Deadweight is defined as elements of the aided activity that would have gone ahead anyway, without assistance. Substitution is where a firm substitutes an aided activity for an unaided activity. Displacement is any reduction in non-aided activities that was generated as a side effect of the policy, through its effect on local product or labour markets, for example. Leakage is the proportion of the outputs or outcomes that occur outside the targeted geographical area or population group. Crowding out is the UK-wide impacts, thought to be imposed through the government's budget constraint. Multiplier effects are the indirect and induced effects generated by the change in intermediate and consumption demand that the policy had produced.

Sometimes evaluations could not report economic impact and instead a 'watered down' version of evaluation was undertaken, assessing the performance of programmes/projects against the original aims, objectives and targets as set out in the business case. Where appropriate, net economic impact (taking into account 'additionality' factors and persistence) was expected to be reported via the

⁷⁰ HM Treasury (2003b).

⁷¹ Single Programme Appraisal Guidance (DTI 2003).

⁷² Assessing the Impacts of Spatial Interventions: Regeneration Renewal and Regional Development (ODPM 2003).

presentation of a GVA-to-cost ratio. However, unlike NICE Health Technology Assessment processes (Chapter 4), there was an unspecified exchequer GVA-to-cost ratio threshold for resource allocation. Given that the RDAs did not conduct economic appraisal through competitive funding rounds, resources were generally allocated on a case-by-case basis.

Of particular interest has been the development of benchmarks or 'ready reckoners'. In a report commissioned by BIS (2009c), statistics were drawn from 280 evaluations of projects and programmes carried out across the UK. Benchmarks were then developed for the key components of 'additionality' for different intervention types according to IEF sub-type. These benchmarks were intended to be used to inform future project appraisal (BIS 2009d) and had the potential to be used as a means of bridging the gap between research and decision making. Indeed, McVittie (2005, p9) discusses that benchmarks can act as "quasi-experimental counterfactuals." However, the benchmarks were highly dependent on the (variable) quality and consistency of the evidence base. Questions could also be raised about issues of generalisability given differing contexts (McVittie 2005). In addition, small sample sizes by IEF sub-group undermined the robustness of the benchmarks and there was no discussion in the report (BIS 2009c) of whether or how the benchmarks would be systematically updated over time when new evidence became available. Overall then, it could be contended that there was not a clear path for evaluation to 'feedback' into economic appraisal processes to inform resource allocation. Cook et al. (2008, p3) note that evaluation was more often used for "exemplification" i.e. providing examples of good practice and lessons learned.

Sensitivity analysis

Sensitivity analysis has been an under-developed area within regional policy evaluation compared to health policy. The HM Treasury's Green Book (2003, p32-33) stated the need to undertake sensitivity analysis at appraisal "to test the vulnerability of options to future uncertainties." However, it does not go into prescriptive detail and instead acknowledges that "expert advice is required to ensure [the techniques] are properly applied" (HM Treasury 2003, p33). The IEF+ (BIS 2009, p22) noted the need to report the confidence level and confidence interval/margin of error achieved for survey

results (if using a probability sample). However, these references to sensitivity analysis are cursory, without specific guidance on how this data should be obtained or applied to decision making.

5.4 Other factors besides evidence which affect the way policy is made

There has tended to be a focus in the regional policy literature on the limited utilisation and impact of evidence on policy making (Leuw 2004; Bachtler 2011). Polverari and Bachtler (2004, p43) state that “it is generally acknowledged that evaluation does not play a major role in determining the overall direction of policy.” They highlight the complexity of the policy making process and the high degree of path dependency based on institutional, political and economic parameters which determine how regional policy evolves. Broader contextual issues are stated to be generally more important than narrow evaluation evidence in determining future policy developments. Some scholars have claimed that evaluations can be commissioned to defend policy decisions rather than inform them (Nilsson et al. 2008). Ferry and Bachtler (2013) also argue that the abolition of the regional tier of government in the UK, described as an example of ‘policy termination,’ reflected the limited role of evidence influencing policy direction at the national level. They explain that austerity measures and the ideological perspective of the incoming Coalition government outweighed evaluations of policy effectiveness. Overall this literature aligns with the tendency in the wider EBPM/KTE literature to focus on the lack of evidence use and the way evidence has been promoted to justify policy decisions. This emphasises the political nature of decision making (Chapter 2).

Generally, however, analysis has not been taken to the next level, which would include distinct exploration of the balance between evidence types and other factors besides evidence which affect the way policy is made, akin to the EBM triad presented in Chapter 4. How evidence is actually incorporated and used in regional policy making processes therefore remains unclear.⁷³

⁷³ This provides a rationale for the empirical analysis undertaken in Chapter 8.

5.5 The RDA national impact evaluation

In 2007, BERR⁷⁴ and the RDAs appointed PricewaterhouseCoopers LLP (PwC) to provide an independent assessment of the impact of spending for each of the nine RDAs and for the RDA network as a whole (PWC 2009a, PWC 2009b). This exercise provides an example of EBRPM in practice, as well as a means of exploring the translation of the IEF guidance into RDA evaluation work. This exercise highlighted that fundamentally the quality and consistency of evaluation studies varied greatly and it was not possible to aggregate the data to generate a robust analysis of the overall economic impact of the RDAs on the national economy as a whole (PWC 2009a; PWC 2009b).

However, the approach used for the RDA national impact evaluation has been critiqued (ONS 2011; Chadwick, Tyler and Warnock 2013). The main issue, highlighted by this extended quotation, is that only “one major proxy of impact [was used] to derive the GVA figures... namely jobs created or safeguarded. Put simply, this was done by assessing the net additional jobs delivered by each intervention at the regional level and multiplying this by an average regional GVA per worker benchmark figure” (Chadwick, Tyler and Warnock 2013, p845-846). “Cost Benefit Analysis then compared how much additional GVA a particular intervention had provided against a given level of public expenditure to derive cost benefit ratios for the different IEF themes” (Chadwick, Tyler and Warnock 2013, p845). Other methodological issues have been identified. Firstly, some evaluations had been unable to apply the IEF and provide estimates of GVA impact (ONS 2011). Secondly, Chadwick, Tyler and Warnock (2013, p845) note that it was not possible to translate all of the core RDA outputs collected⁷⁵ into GVA estimates “within the time and budget constraints of the RDAs’ evaluation plan.” Thirdly, overall the number and relative simplicity of methods used to estimate GVA was problematic. In particular, the average regional GVA per worker figures (and economic impact multipliers) were derived from official ONS estimates, and did not reflect the diversity of the regions or sectors (ONS 2011). Finally, the range of RDA

⁷⁴ The Department for Business, Enterprise and Regulatory Reform (BERR) merged with the Department for Innovation, Universities and Skills (DIUS) to become the Department for Business, Innovation & Skills (BIS) on 5 June 2009.

⁷⁵ As set out in the RDA tasking framework (cited in DTI 2006, p141-143). Core outputs included: “jobs created or safeguarded; people assisted into employment; skills assists; businesses created; businesses assisted; and land remediated.”

interventions meant that estimates of impacts were not directly comparable across IEF themes. What impact the (more prescriptive) IEF+ guidance may have had on the evaluation evidence base is unknown given the abolition of the RDAs and dismissal of the IEF framework.

It could also be argued that the RDA national impact evaluation fell short compared to the large-scale national evaluations of Business Link (Mole et al. 2008), the New Deal for Communities Programme (CLG 2010a; 2010b), and Sure Start (Belsky, Barnes and Melhuish 2007). Intriguingly, certain key aspects of regional policy evaluation at a national level have thus tended to be overlooked. For instance, while decentralisation of funding and powers to the regional and then local levels have been a key focus of New Labour and then Coalition policy, a study has not yet been undertaken to assess whether such institutions and policies have (or will) achieve superior outcomes to a more centralised approach and to examine the national efficiency of regional policy.

Overall, the RDA national evaluation report (PWC 2009a) did actually produce broadly positive conclusions on the effectiveness and efficiency of the RDAs. However, Ferry and Bachtler (2013, p269) note that “these were ignored and indeed contradicted by the political narrative in the debate on [policy] termination”, particularly given the context of economic crisis and austerity. Instead, there was a focus on the problems faced by the RDAs over accountability and value for money issues in policy discourse (NAO 2010; Great Britain, Parliament, House of Commons 2010).

5.6 Conclusion

This chapter has contributed in a number of ways to address the first research question. It was found that there has not been a formal ‘evidence based movement’ within regional policy equivalent to EBM within health policy. Partly because of this, less attention has been paid in the literature to evidence hierarchies, the role of research credibility and the incorporation of evidence into decision making. This analysis identified that evaluation evidence has not played a major role in policy development and has shown that there is a gap between practice and academic thinking.

This chapter has also demonstrated, however, that the context of regional policy is very different, and in some ways is more complex than health. The complexities

associated with regional policy's multifaceted agenda, structures and mechanisms alongside its inherently political character transform the nature of the evaluation process. Regional policy evaluation has very different aims from evaluations conducted within the EBM paradigm and needs to address political issues and choices alongside intervention efficacy. This provides the foundation for the following chapter where the epistemological and applicability implications of extending an EBM approach to regional policy evaluation are analysed.

Chapter 6

Extending an EBM Approach to Regional Policy Evaluation

6.1 Introduction

This final literature review chapter builds upon the groundwork so far provided and addresses the first research question. The overarching EBPM debates have been explored focusing on case studies of health policy and EBM (Chapter 4) and regional policy and impact evaluation (Chapter 5). By using a mirrored analysis in this way it has been possible to highlight key differences across the sectors in terms of generating and using evaluation evidence for investment prioritisation. In this chapter a direct comparison is made across the sectors and the epistemological and applicability implications of extending an EBM approach to the regional policy evaluation context are surfaced.

6.1.1 Theoretical background

This comparative literature review predominantly focuses on debates around evidence types and the role of research credibility. Therefore the unabridged discussion of the theoretical background for this chapter is provided in Chapter 2 (predominantly in section 2.3).

It was found in Chapter 1 that a significant underlying debate for regional policy evaluation, which has certainly come to the fore recently, has focused on trust in the reliability of research findings. This has led to a call for the extension of approaches more closely aligned with EBM to wider social policy such as the use of RCTs, the establishment of a 'NICE for social policy' and the use of quasi-experimental approaches and economic evaluation (Haynes et al. 2012; Cabinet Office 2013; BIS 2014b). Caution is needed, however, as EBPM is a contested concept⁷⁶ across policy domains, and within social science (Chapter 2) and Evidence Based *Regional* Policy Making (EBRPM) in England is an under-researched area (Chapter 1). Although the academic literature is well established and becoming increasingly sophisticated, the practice of regional policy evaluation in England by the institutions charged with its

⁷⁶ as are elements of EBM, as shown in Chapter 4.

implementation has not kept pace with this development. Less attention has been paid to the evaluation of UK regional policy instruments (i.e. non EU policy instruments) and to the processes of undertaking regional policy evaluation in practice (Chapter 5). Thus, given the recent shift of focus for regional/local policy evaluation within the wider 'what works' agenda, the implications of extending an EBM approach are relatively unknown.

As a (previously) practising RDA evaluation officer and research economist within the health sector, I had some prior understanding of the issues associated with extending an EBM approach to regional policy evaluation. Reflecting upon the above findings and drawing upon this experiential knowledge, it could be hypothesised that undertaking a comparative review across the health and regional policy sectors (cases) would highlight that parallel EBM processes for evaluation and investment prioritisation would be played out in different ways within each context (Skocpol and Somers 1980). This is examined in the sections that follow.

6.2 A comparison of EBM and regional policy evaluation

When undertaking a comparison of EBM and regional policy evaluation, it is necessary to reflect on the factors that might influence the adoption of an evidence based approach. To analyse EBPM debates across policy domains, a conceptual framework was developed to structure examination of the existing literature and theories (Chapter 1). In particular, a paper by Sutcliffe and Court (2005) was drawn upon which distinguished between three central theoretical questions in the EBPM literature: what kinds of evidence are used and the role of research credibility ('what'); the issues surrounding the way in which evidence is incorporated into the policy making process ('how'); and what are the other factors besides evidence which affect the way policy is made ('other factors'). These three cross-cutting debates were used to draw parallels across the health and regional policy sectors in Chapters 4 and 5 (with Chapter 5 focusing on the timeframe of the RDAs).

The analysis undertaken in the preceding chapters is summarised in Table 11 below. Points made in square brackets indicate a shift in thinking under the Coalition government. A simplified discussion is to follow presenting two ends of the EBPM spectrum, with EBM at one end and regional policy evaluation at the other.

Table 11: A comparison of EBPM debates across EBM and regional policy evaluation

	EBM	Regional policy evaluation
What	<ul style="list-style-type: none"> • Explicit evidence hierarchies • RCTs/systematic review • Economic evaluation • Clinical expertise 	<ul style="list-style-type: none"> • Implicit evidence hierarchies • Method Based Evaluation • Hybrid independent expert/participatory approach • Impact evaluation • [Inclusion of quasi-experimental methods]
How	<ul style="list-style-type: none"> • NICE • [Extension to wider health & social policy] 	<ul style="list-style-type: none"> • Division between central and regional agencies • National RDA impact evaluation • [Establishment of the 'What Works Centre for Local Economic Growth']
Other factors	<ul style="list-style-type: none"> • Patient preferences • Politics • Phronesis 	<ul style="list-style-type: none"> • Politics • Phronesis

What

The factors identified within Table 11 are closely interlinked. Focusing on 'what' evidence, the underlying consideration cutting across EBM and Evidence Based Regional Policy Making (EBRPM), has been the pursuit of internal validity by identifying credible counterfactuals within study designs (i.e. to guarantee that the outcome seen is due to the intervention). To consider the analysis undertaken in Chapters 4 and 5, it was found that explicit evidence hierarchies are an integral component of the EBM approach, with RCTs and systematic reviews placed at the top and case study reports usually at the bottom of the hierarchy. In contrast, within Method Based Evaluation (MBE) in regional policy, evaluation guidance has shaped an implicit evidence hierarchy whereby macro, quantitative approaches (such as econometric modelling) are placed at the top, and micro, more qualitative approaches (such as case studies) are placed at the bottom. Over time, a *hybrid* RDA evaluation model developed, with the intention of integrating various forms of evidence and processes of collating

evidence from independent expert through to action learning approaches. However, “a commitment to a participatory approach... (which involves service users, practitioners and evaluators working together) emphasises research designs that would typically score low on such scales” (Nutley, Powell and Davies 2013, p12). A focus on internal validity has come to the fore again under the Coalition government, with a renewed call to establish control groups within study designs and the promotion of quasi-experimental methods.

However, Cartwright and Hardie (2012) argue that internal and external validity are confused in the assumptions underlying evidence hierarchies. Contandriopoulos et al. (2010, p457) also argue that evidence is heavily context-dependent and that relevance to policy-makers “has less to do with internal validity [i.e. scientific rigour] than with external validity” (i.e. generalisability and the perceived alignment with existing knowledge). This leads on to the debate between methods-based versus problem-based research in regional science more generally (Markusen 2015). Markusen (2015, p8) argues that an excessive focus on methods-driven research and practice can hamper the development of appropriate policy solutions. Sefton (2000, p21-22) makes a link between overly method-driven study designs and ‘reductionist’ “run of the mill” evaluation outputs. Therefore, evidence hierarchies have been regarded as inappropriate when considering the role of ‘policy relevant evidence’ as they do not consider the applicability of findings to policy concerns (Abeyasinghe and Parkhurst 2013). In turn, if EBPM is to be based upon a foundation of such evidence hierarchies, then policy making will be biased towards those issues conducive to the methods promoted by them. Immediate short term interventions will become the focus, rather than long term structural or social changes where it is difficult or impossible to implement ‘rigorous’ study designs (Barnes and Parkhurst 2014). This highlights the fundamentally political nature of evidence.

Closely related to the above debates is ‘what’ evidence is considered credible (for resource allocation in particular). The underlying consideration cutting across EBM and EBRPM has been the pursuit of economic evaluation evidence for such decision making. To consider the analysis undertaken in Chapters 4 and 5, it was found that within EBM the NICE process has led to demand for evaluations based upon the use of cost-effectiveness modelling. In contrast, within regional policy it has been

acknowledged that monetising all costs and benefits is impracticable and so impact evaluation has been the method detailed in RDA evaluation guidance.

How

‘How’ evidence is incorporated into the policy making process is therefore intrinsically linked to the above discussion. The underlying consideration cutting across EBM and EBRPM has been the institutional framework (or lack of it) demanding evidence centrally, peer-reviewing it and potentially using it to inform decisions. It was found that the establishment of NICE effectively institutionalised EBM within the health sector, and that the Health Technology Appraisal (HTA) process aims to incorporate both expert clinical and academic advice as well as patient opinion. Incremental Cost-Effectiveness Ratios (ICERs) are compared against a specified NICE threshold for resource allocation decisions. The NICE HTA process essentially generates evidence of pilot tested policy interventions, and such evidence can be fed directly into the requirement for NHS bodies (such as Clinical Commissioning Groups) to fund medicines/new technologies that NICE has endorsed. Therefore NICE, and by implication EBM, has a well-defined but relatively limited role (albeit contested at the margins) with respect to the whole of the NHS. In contrast, for regional policy (and health and social care policy more widely) the potential scope for EBPM is greater and could include all project/programme activity. Yet there has not been an institutional framework such as NICE to incorporate evaluation evidence into regional policy making, and the division between central and regional agencies has been somewhat unclear. The RDA national evaluation was the first attempt to provide an independent assessment of the *net* impact and economic value of RDA interventions (Chadwick, Tyler and Warnock 2013)⁷⁷.

An important factor to consider is the relative scale of activity across the sectors, and this can be reflected by the level of government expenditure on the NHS and the RDAs. For instance, in 2009/2010 total government expenditure was £669.26bn, of which a large proportion (£100.2bn) was spent on the NHS (Chapter 4) and a small proportion (£2.26bn) was allocated to the RDAs’ combined single pot budgets (Chapter 5). This is likely to be a key factor that would influence the adoption of an evidence based

⁷⁷ See later discussion on the WWG in section 6.4.3.

approach and the need to pilot test policy instruments and delivery. Although difficult resource allocation, rationing, and priority setting questions are raised for both sectors, the larger proportion of the Government's budget spent on health clearly places pressure on the NHS budget to be spent "wisely, fairly and transparently to secure the best possible outcomes for both patients and the taxpayer" (NHS ENGLAND 2015, p50).

Other factors

In terms of 'other factors' besides evidence which influence decision making, the analysis undertaken in Chapters 4 and 5 found that within EBM the role of clinical expertise, external evidence and the role of patient preferences has been explicitly considered in the literature. In contrast, distinct exploration of the role of experiential and expert knowledge and the balance between evidence types and other factors which affect the way policy is made (akin to the EBM triad presented in Chapter 4) was found to be lacking in the regional policy literature.⁷⁸

It was found that there are two underlying, interwoven, considerations cutting across EBM and EBRPM. The first focuses on the perceived role of politics within EBPM, and the second focuses on the role of individual decision making processes. It was generally acknowledged across both sectors that there are other, external, factors besides the use of evidence which are important to policy making. It is somewhat of a paradox then that the underlying assumption in much of the EBPM literature (Chapter 2) is that the influence of evidence should be elevated regardless of these other factors. This is a key finding of the literature review undertaken by Rutter, Hawkins and Parkhurst (2013) who identify a deterministic focus in the literature of 'getting evidence into policy'. The conceptualisation that policy should be based on evidence implies that politicisation of the process should be reduced and disregards the political nature of evidence.

⁷⁸ This provides a rationale for the empirical analysis undertaken in Chapter 8 to explore the generation, communication and use of evaluation within regional policy organisations (the RDAs) to reveal how evidence was incorporated into policy making processes and the role of other factors besides evidence.

Secondly, how precisely different types of evidence and considerations of context and resources are combined cognitively by individuals alongside their inherent human 'practices' and standards such as fairness, truthfulness, trust and honesty (termed 'Phronesis' by Sanderson 2003), is a question that remains to be addressed academically and within policy. These external and internal 'other factors' have profound implications for the producers of evidence within EBPM. All of these reflections are expressed within an extended quotation by Rutter, Hawkins and Parkhurst (2013, p 28) who conclude that:

Democratically representative decision makers can be informed by multiple bodies of evidence, yet they still can apply social values to judge between the different outcomes that acting on various evidence bases will achieve. However, greater appreciation of the political nature of decision making does not mean that the concerns with quality of evidence and unbiased reviews of evidence become irrelevant. Indeed, a focus on the appropriate use of evidence (instead of simply use) could potentially allow the Knowledge Transfer and Exchange field to move forward.

Overall, the above discussion highlights key differences across the health and regional policy contexts. It also reveals assumptions embedded within EBM and within EBRPM which have direct implications for an extension of an EBM approach to regional policy (and to wider health and social policy). Mirroring such analysis across both sectors also surfaced further theoretical, methodological and practical considerations focused primarily on the types of evidence that are produced ('what evidence'). These are to be discussed further in the following sections.

6.3 Epistemological challenges

Hoffman et al. (2012, p21) note that different epistemological perspectives can lead to disagreements concerning "the nature of knowledge and how it is discovered or co-created." Such philosophical differences between EBM and wider social policy research have been studied in the literature (Sefton 2000; Sefton et al. 2002; Sefton 2003; Somekh et al. 2005). Applying such analyses to regional policy evaluation specifically, the following characteristics are of significance: study populations; timing and purpose of the evaluation; data type; understanding impact; and philosophical perspective. Drawing upon the analysis undertaken in the preceding chapters and applying these identified characteristics, an epistemological comparison of EBM and regional policy evaluation can be made (Table 12). Points made in square brackets indicate a shift in

thinking under the Coalition government. Once again, it should be borne in mind that the analysis in Table 12 is relatively high-level and presents two ends of the EBPM spectrum.

Table 12: An epistemological comparison of EBM and regional policy evaluation

Characteristics	EBM	Regional policy evaluation
Theoretical framework	<ul style="list-style-type: none"> • Biomedical model of health 	<ul style="list-style-type: none"> • Spatially unbalanced growth • [People based policy]
Study populations	<ul style="list-style-type: none"> • Individuals 	<ul style="list-style-type: none"> • Institutions • Places
Timing	<ul style="list-style-type: none"> • Ex-ante 	<ul style="list-style-type: none"> • Ongoing • Ex-post
Purpose	<ul style="list-style-type: none"> • Effectiveness 	<ul style="list-style-type: none"> • Process • Effectiveness
Data type	<ul style="list-style-type: none"> • Quantitative 	<ul style="list-style-type: none"> • Qualitative • Quantitative
Understanding impact	<ul style="list-style-type: none"> • Understanding 'what' 	<ul style="list-style-type: none"> • Understanding 'how', 'why' and 'where' • Understanding 'what'
Philosophical perspective	<ul style="list-style-type: none"> • Positivism 	<ul style="list-style-type: none"> • Plurality of perspectives

Source: adapted from Sefton et al. 2002, p 27-29

The factors identified within Table 12 are closely interlinked and some have been touched upon elsewhere in the discussion. The main point to be made is that there is a fundamental difference in the theoretical frameworks of EBM and regional policy evaluation. EBM, with a biomedical model at its core, focuses on individuals' biological outcomes to medical intervention (Somekh et al. 2005). In regional policy, however, it is necessary to recognise "the interactive relationship between individuals and their environment" (Sefton et al. 2002, p28). Sefton et al. (2002, p28) argue that "conventional approaches to economic evaluation will be less suited to evaluating programmes that have the community, rather than the individual, as the focus of interest."

As such the timing, purpose and methodological approach to evaluation differ across EBM and regional policy. Evaluations to determine effectiveness are mostly carried out ex-ante for EBM (i.e. as pilots or trials) compared to ex-post evaluations for regional policy. On-going or mid-term evaluations may also be carried out within regional policy to review process issues. Within EBM, Zwarenstein and Treweek (2009, p998) observe

that the vast majority of clinical trials focus squarely on the question of ‘what works’ and are “explanatory” (i.e. designed to test a hypothesis in a highly controlled context), rather than “pragmatic” (i.e. designed to identify interventions that might produce beneficial outcomes in practice). They note that fewer than 100 ‘pragmatic’ designed RCTs have been identified out of the 250,000 clinical trials listed by the US National Library of Medicine. Nutley, Powell and Davies (2013, p15) point out that “those interested in evidence-based practice also want answers to other questions besides what works, such as what matters and what is acceptable.”

This leads on to a long-standing philosophical debate termed the ‘causal wars’ (Scriven 1994) with ‘positivism’ at one end of the spectrum and ‘constructivism’ at the other. The analysis undertaken in Chapters 4 and 5 identified that EBM is aligned with a positivist view of the world, as is MBE in regional policy evaluation. “With perfect information, appraisal, evaluation and optimal policy choice become purely technical problems” (McVittie and Swales 2003, p6). Scriven (1994) used the term ‘black box’ for this type of evaluation, due to the focus on outcomes, “with no explanation or understanding required with regard to how recorded outcomes might have been produced” (Salter and Kothari 2014, p2). However, Armstrong and Wells (2006b) identify that Structural Funds evaluation practice have also drawn upon realism and constructivism. Constructivists argue that “theories and realities are not ‘out there’ waiting to be discovered or uncovered, but are constructed in the minds of individuals or in the discourses of groups” (Kushner 1996, p189). Kushner (1996) critically analyses constructivism as it has appeared in the field of evaluation and presents it as an overreaction to the problems of objective reality. Somewhere in the middle of the spectrum are realist, theory-based approaches. These are ‘explanation-driven’ and aim to uncover what works, for whom, and under what conditions (Pawson and Tilley 1997), and aim to be responsive to changes in context and knowledge (Van der Knaap 2006). Therefore, it could be argued that overall the RDA evaluation model drew upon a plurality of theoretical perspectives. In turn, this was reflected in the aim to integrate various forms of evidence and processes of collating evidence (Chapter 4, Figure 15).

Such theoretical differences in EBM and regional policy have been reflected by the methods which have been considered appropriate for application given the differing

contexts. Within EBM there has been a focus on the use of RCTs, whereas RCTs have not been widely applied within the regional policy sector.⁷⁹ In the following sections, the methodological and practical challenges to promoting RCTs within regional policy evaluation are therefore further analysed.

6.4 Applicability challenges

6.4.1 Experimental methods

Sefton (2000; 2003) argues that problems of measurement, attribution, and interpretation are more acute within social policy research. To undertake an experimental study design, in particular a RCT, Canter (2012) presents a number of assumptions that need to be met. A distinct causal variable needs to be identified (the independent variable, IV). Clear, expected effects need to be specified and measured (the dependent variable, DV). The main influences on the DV beside the IV need to be determined so that an appropriate 'control' group can be identified. The interactions between IV's ideally need to be relatively straightforward and not recursive or contingent. For an RCT, entities need to be randomly assigned to conditions in which the IV is present or in which it is not. Reflecting upon this, it could be argued that such RCT approaches within EBM, and traditional welfare approaches within regional policy evaluation, "imply perfect knowledge about policy objectives and the way in which policy operates" (McVittie and Swales 2003, p6). Drawing upon the analysis undertaken in the preceding chapters and applying these identified assumptions, a comparison of EBM and regional policy evaluation can be made (Table 13). Once again, the analysis in Table 13 is relatively high-level and presents two ends of the EBPM spectrum.

⁷⁹ For instance, the 'What Works Centre for Local Economic Growth' systematic review of employment training reported on only 2 RCTs (WWG 2014, p19).

Table 13: A methodological comparison of EBM and regional policy evaluation

Characteristics		EBM	Regional policy evaluation
Baseline dependent variable (DV)	Rationale	Homogenous rationale for intervention	Heterogeneous rationale for intervention
Causal variable(s) (IV)	Intervention	Well-defined intervention	Poorly-defined intervention
		Single intervention	Multiple intervention
		Static intervention	Non-static intervention
Change in the DV	Outcomes	Well-defined objectives	Poorly-defined objectives
		Single outcomes	Multiple outcomes
		Quantitative outcome measures	Quantitative & qualitative outcome measures
		Short-term outcomes	Long-term outcomes
		High level effects	Low level effects
Factors which influence DV apart from IVs	Study population	Individual-clinical level	Population-policy level
		Homogenous treatment population	Heterogeneous beneficiary population
	Delivery agent	Single agency delivery	Multi-agency delivery
			Multi-level delivery
	Socio-economic and spatial context	Independent of context	Highly dependent on context
Interactions between IVs	Cause-effect model	Linear cause-effect models	Nonlinear cause-effect models

Table 13 demonstrates that RCTs and experimental approaches are more reconcilable to EBM and biomedicine than to regional policy evaluation as there is more likely to be a homogenous medical condition of study, homogenous patient population, well-defined intervention, and intermediate and final endpoint data collected on specific, large, short-term, measurable outcomes to ascertain intervention effectiveness. The use of a control group through an RCT can hold all other factors constant (*ceteris paribus*) to identify intervention effectiveness. This is not to say that health outcomes research and cost-effectiveness modelling is a simple practice. Highly complex models

requiring a great amount of technical expertise are required. However, a number of complexities associated with regional policy evaluation have been highlighted in the literature which forms the basis of comparison to EBM in Table 13. These are discussed in the following paragraphs.

The factors identified in Table 13 are intrinsically linked. In terms of the rationale for intervention, Diez (2002) has written about the 'systemic nature' of regional policy whereby myriad interventions may be targeted at multiple beneficiaries, including companies, institutions, communities and areas, potentially all with differing needs for intervention. Diez (2002) has also written about 'dynamism and flexibility', identifying that interventions are fluid and multi-faceted in nature and may change over time according to changing socio-economic or political conditions. Likewise interventions may be refined due to feedback from evaluation or monitoring data (European Commission 2007). Polverari and Bachtler (2004) agree that often interventions are poorly defined with 'intangible' and often changing policy objectives. Such objectives may be 'conflicting' (Bachtler 2001) and Stern (2003) highlights that programmes of interventions often 'overlap'. McVittie and Swales (2003) take this a step further noting that policy objectives may be purposefully kept 'vague'.

Scholars have also noted that multiple outcome measures can be synonymous with regional policy interventions (Armstrong 2000), which makes it more difficult to make comparisons between programmes unless one performs better on all counts (Sefton et al. 2002). In addition, the outcomes of some regional policy interventions are identified to be qualitative by their very nature and not very amenable to measurement, such as those focused on Strategic Added Value and on wider socio-economic and environmental impacts beyond GVA (DTI 2006). 'Time lags' (Polverari and Bachtler 2004) and 'extended timescales' (Stern 2003) for undertaking evaluation are identified as being important, highlighting the long term nature of regional policy. Thus it has been acknowledged that effects may be hard to detect, and this may be compounded by the low level effects of many regional policy interventions whereby impact may be small relative to the scale of the market failure (Sefton et al. 2002; Rhodes, Tyler and Brenan 2005).

It is widely recognised that broader contextual issues are important in regional policy evaluation. Diez (2002) has written about the role of 'embeddedness' highlighting that intervention efficacy is highly dependent on institutional, political and economic parameters and on the complexity of the policy making process. Polverari and Bachtler (2004) note that the negotiated nature of regional policy means there is a need to involve a wide range of actors in the policy process (partners, stakeholders, beneficiaries, etc.). Multiple agencies and every level of government may be involved in regional policy making and delivery, from EU institutions to national governments, regional agencies and local authorities (Bachtler 2001). Policy may involve actors in the public, private and voluntary sectors (Bachtler 2001). Diez (2002) explains that local and geographical variations in intervention effectiveness occur as there is local autonomy, co-operation and partnership in the delivery of many interventions. As such, "complex interactions are produced in multiple areas and at different levels of effects" (Diez 2002, p290).

It is acknowledged that in regional policy, evaluations are conducted in a 'real life' setting where the breadth and complexity of practice means it is often impracticable to isolate the effects of a particular intervention. It is also generally agreed that "the method of randomised experimentation tells us nothing about whether the same results would be seen elsewhere, or would work in a different policy environment" (Rutter, Hawkins and Parkhurst 2013, p27). However, the overall cause-effect model is less commonly considered in the literature. Stern (2003) points to 'uncertain implementation chains.' This could suggest recursive and/or contingent interactions between resources, activities, results, effects and impact. It has therefore been argued that the use of controls in natural (social policy) settings is paradoxical as the 'confounding' variables are likely to be relevant to the processes under study (Canter 2012).

In contrast, theory-based approaches aim to build understanding of the reasons for effectiveness and the circumstances under which results are likely to be replicated, thus going some way to account for the complexity of the context and contingency between variables. For instance, the use of logic models involves mapping out the causal chain from inputs to outcomes and impact, and testing the underlying assumptions (Tavistock Institute 2003). Similarly, 'Realistic Evaluation' programme

theory defines a set of explicit and testable assumptions or hypotheses about how a programme is supposed to achieve its goals. This is then tested through a 'Context-Mechanism-Outcome' configuration to understand how a specified intervention will produce given outcomes in certain contexts (Pawson and Tilley 1997). The purpose of clearly specifying the research context and any assumptions are to enhance the 'transferability' of the research findings to other settings (Guba and Lincoln 1994). These approaches draw upon both quantitative and qualitative data to understand whether and how the intervention may be replicated in other settings. It is imperative that data are collected to test the underlying assumptions of the causal links (White 2009).

Davis (2005, p275) identifies issues with the operationalisation of the concept 'context matters.' He notes that although realistic evaluation has been applied within social policy settings, these have been discrete projects with specific initiatives and have thus been limited in the scope of the systems evaluated. Stufflebeam and Shinkfield warn that evaluators using this approach may "focus attention on theory developed early in the program and later discover that the program has evolved to be a quite different enterprise from what was theorised at the outset" (2007, p187). However, of importance is that a way forward, incorporating analysis of both the effectiveness of the intervention and implementation, has been identified.

Sefton (2000) argues 'Theory of Change' models and standard economic approaches to evaluation both follow an input-outcome framework. Accordingly it is theoretically possible to examine both variations in costs and outcomes as well as taking into account context variables (Sefton 2000). It could be envisaged that this would require economic models of both the relationship between policy intervention and the outcome metrics and the relationship between socio-economic metrics and the outcome metrics. Weiss, Bloom and Brock (2014, p778) agree that evaluating program implementation and estimating program effects can be integrated into the same study, which may "help to identify factors that lead to variation in program effects and thereby support more systematic data collection." They have developed a conceptual framework for studying the sources of variation in program effects to identify conditions and practices that are associated with larger and more positive effects, explicitly accounting for context (Weiss, Bloom and Brock 2014).

Overall, the analysis in Tables 12 and 13 highlights that there are considerable theoretical and methodological challenges with applying RCTs within regional policy. There are also a number of practical difficulties. RCTs have also been deemed unfeasible on cost and ethical grounds within social policy. Thus, the inadequacy of research funding levels for experimental study designs has been stressed previously in regional policy evaluation guidance (DTI 2006, p46) and 'proportionality' has been a strong element of UK (HM Treasury 2003b) and EU guidance (European Commission 2007). Ipsos-Mori (2012) argue that the inclusion of a control group of equal size to a treatment group in an evaluation will normally double the recruitment and fieldwork costs associated with monitoring outcomes. In addition, ethical issues for applying RCT methodology to social policy are also widely documented. Nutley, Powell and Davies (2013, p11) point out that it may be difficult or impossible to implement 'blinded' RCT designs "that ensure that individuals, practitioners and analysts are unaware of whether subjects are in experimental or control groups" to avoid a 'placebo' effect. This applies to individuals but also to other units of randomisation such as local authorities and firms (Ipsos-Mori 2012). Sefton et al. (2002) add that even if a good initial match can be made between 'experimental' and 'control' areas, it is unlikely that, in reality, circumstances would have remained similar throughout the evaluation period.

6.5 Recent developments for regional policy evaluation in 2015

6.4.2 Quasi-experimental methods

There has been a recent shift in focus in the political narrative towards the promotion of quasi-experimental comparison group designs (BIS 2014b). The academic literature on quasi-experimental methods has become increasingly sophisticated (Isserman and Rephann 1995). However, in a review of the use of quasi-experimental methods in regional research, Feser (2013, p44) reports:

There is still progress to be made in improving matching methods, making more extensive use of time-series designs, undertaking more systematic sensitivity testing and checks for the robustness of findings, focusing greater attention on effect heterogeneity.

Pirog (2014) notes that econometric approaches could be a powerful tool to address selection bias issues within public policy research. However, she emphasises that a key

issue undermining quasi-experimental research designs and the use of statistical/econometric approaches is the co-dependency with data availability and quality. Pirog (2014) suggests a need for greater data-linking across agencies and programmes and to geospatial data. Pirog (2014, p537) argues that “in the absence of better econometrics, researchers will continue the movement into experimental research.” Likewise, Isserman and Rephann (1995) draw attention to the need for longitudinal data given that regional policy is a long term endeavour and results may not be seen for 25 years. In addition, the need for a clear strategy for regional policy is identified to inform the development of a spatial comparator. Isserman and Rephann (1995) discuss regional policy clearly in terms of supporting disadvantaged places/regions. However, in Chapter 1 it was identified that for UK regional policy there has been a balance between potentially conflicting objectives for rebalancing the economy and sustaining national competitiveness.

To focus on the recent Regional Growth Fund (RGF) evaluation scoping study as a case in point, a quasi-experimental approach has been proposed and it is reported that matching will be the technique used for identifying a counterfactual (BIS 2014b, p16) i.e. the use of propensity score matching, difference-in-differences and/or fixed effects modelling. In reality, there are a number of issues involved with taking such an approach to the RGF evaluation, and it is ultimately dependent on the data quality and availability (particularly beneficiary data availability). The design of the RGF poses several challenges to the development of an appropriate counterfactual. It has not been designed as an area based initiative and therefore has no rigidly defined spatial boundaries for the programme.⁸⁰ This is despite one of its core objectives being to rebalance the economy and create sustainable private sector jobs in areas highly dependent on public sector employment. Therefore, the scale of any displacement, multiplier effects and crowding out are unlikely to be understood at a national level to provide estimates of the net additional economic impacts of the RGF.⁸¹ A mixed methods approach has therefore been proposed, including the use of methods more

⁸⁰ Rhodes, Tyler and Brennan (2005) have previously identified the issue of a mismatch in the boundaries of area-based initiatives and statistical units of data collection.

⁸¹ Similarly, Wilson (2013) has identified that for the New Deals for Communities evaluation, individual level effects, rather than national levels effects, may be identified suggesting the issue may be the level at which analysis takes place, rather than the use of a quasi-experimental design per se.

closely aligned with those traditionally used by the RDAs (i.e. case study analysis and beneficiary surveys).

6.4.3 NICE for social policy

Another extension of EBM under the Coalition government has been the (partial) establishment of an institutional framework which critically analyses the nature of the evidence base and shapes ‘what’ evidence types are demanded and deemed credible.⁸² A ‘What Works Centre for Local Economic Growth’ (WWG) has been established, conceptualised as part of a ‘NICE for social policy.’⁸³ As stated in the Civil Service Reform Plan (HM Government 2013, p 17):

An important element of this is a clear understanding of “what works”, building on evidence from policy in practice... In the same way that... NICE advises the NHS, the Cabinet Office will review the value of creating a similar institute that can test and trial approaches and assess what works in major social policy areas, so that commissioners in central or local government do not waste time and money on programmes that are unlikely to offer value for money.

However, it has been argued that there is misunderstanding surrounding the workings of the ‘NICE model’ and strong emphasis for the ‘What Works Centres’ has been placed on the supply, rather than the demand for evidence (Alliance for Useful Evidence 2014). Indeed, the WWG has been tasked with undertaking systematic reviews of current evidence to uncover drivers for local economic growth/employment and to transfer such evidence to policy makers within Local Authorities and Local Enterprise Partnerships (LEPs) (Cabinet Office 2013). The mandate of the WWG is summarised in Table 14.

⁸² In addition, a BIS ‘Expert Peer Review Panel’ was launched in January 2014 “to review all evaluations that make claims about impact or value for money of policy” (BIS 2014f, p12).

⁸³ Referred to in the Open Public Services White Paper (Cabinet Office 2011); Department for Business, Innovation and Skills Innovation and Research Strategy (BIS 2011); Civil Service Reform Plan (HM Government 2012); ‘What works: evidence centres for social policy’ (Cabinet Office 2013).

Table 14: Tasks for the What Works Centre for Local Economic Growth

Task	Sub-task
Generate a summary of evidence synthesis	<ol style="list-style-type: none"> 1. Undertake systematic assessment of relevant evidence and produce a sound, accurate, clear and actionable synthesis of the global evidence base which: <ul style="list-style-type: none"> • assesses and ranks interventions on the basis of effectiveness and cost-effectiveness; • shows where the interventions are applicable • shows the relative cost of interventions • shows the strength of evidence on an agreed scale
Translate the evidence	<ol style="list-style-type: none"> 1. Produce and apply a 'common currency' - a common set of standards in each area for comparing the effectiveness of interventions 2. Put the needs and interests of users at the heart of its work
Share the evidence	<ol style="list-style-type: none"> 1. Publish and share findings in a format that can be understood, interpreted and acted upon
Promote good evidence	<ol style="list-style-type: none"> 1. Identify research and capability gaps and work with partners to fill them 2. Advise those commissioning and undertaking innovative interventions and research projects to ensure that their work can be evaluated effectively

Source: 'What works: evidence centres for social policy' (Cabinet Office 2013, p5)

It could be argued that the WWG's mandate is ambitious. For the evidence synthesis task, the WWG has committed to use an explicit evidence hierarchy to appraise the quality of studies. The hierarchy used has been based upon the Scientific Maryland Scale (SMS) developed by Sherman et al. (1998) within the field of crime statistics. The SMS is a five-point scale ranging from one for studies based on simple cross sectional correlations, to five for randomised control trials (a simplified version is presented in Table 15). Sherman et al. (1998) indicate that confidence in the results is highest at level five and level three should be the minimum level required to achieve reasonably accurate results.

Table 15: The Scientific Maryland Scale

Level	Study design	Summary
Level 5	Randomised Control Trial	Level 5 includes randomisation into treatment and control groups
Level 4	Difference-in-differences study identifying causality <i>ceteris paribus</i>	Level 4 makes a before-after comparison using a control group, but ensures that all other factors are held constant (<i>ceteris paribus</i>) to isolate the effect of the intervention
Level 3	Difference-in-differences study	Level 3 makes a before-after comparison using a control group
Level 2	Before and after study	Level 2 compares the outcomes before and after the study without establishing a counterfactual
Level 1	Simple cross-sectional correlations	Level 1 ascertains the correlation between an intervention and outcomes without establishing a counterfactual

Source: adapted from Sherman et al. (1998, p4-5); What works Centre for Local Economic Growth (2014, p16)

However, as highlighted by the above discussion, vulnerability for the WWG lies in the evidence base it is to synthesise, which includes studies commissioned by the RDAs. The first WWG systematic review was published in April 2014 (WWG 2014) focused on evaluations of training programmes. The study identified quality and consistency issues with the evidence under review. Almost 1000 policy evaluations, evidence reviews and meta-analyses from the UK and other OECD countries were reviewed, but it was found that only 2 reports were categorised at ‘level 5’ on the SMS scale, 11 at ‘level 4’ and 58 at ‘level 3,’ indicating that most studies did not include a control group (WWG 2014, p19). The systematic review was unable to use meta-analysis to provide a pooled estimate measure of effectiveness for comparison between interventions. The authors of the review (WWG 2014, p4) noted that the limited evidence base had constrained the use of such methods in comparison to other policy areas such as medicine and education.

As shown in Table 14, the next phase of the WWG’s mandate is to move beyond synthesising primary studies towards a knowledge transfer role. Interestingly, according to the WWG’s work-plan (RCUK 2014), an online toolkit (decision tool) is planned as part of this phase, based on the research synthesis work and allowing users to compare policies according to a ‘common currency’ (i.e. a common set of standards

in each area for comparing the effectiveness of interventions). The question of how to update such systematic reviews as new evidence is produced has not been addressed in the WWG work-plan (RCUK 2014). Currently, there has been no attempt (that the researcher is aware of) to develop, or update, benchmarks to inform future project appraisal along the lines of the work commissioned previously by BIS (BIS 2009c) for key components of ‘additionality’ for different intervention types.

Therefore, drawing upon the above discussion, although the potential for techniques such as Bayesian meta-analysis⁸⁴ to enable the systematic update of the evidence base may be identified, this would be dependent upon work on a ‘common currency’ or benchmarks to have been undertaken. Bayesian meta-analysis may enable the inclusion of ‘informed priors’ (Spiegelhalter and Best 2003) to statistically combine Return on Investment (ROI) and ‘additionality’ benchmarks with new evidence as it is produced. Techniques such as probabilistic sensitivity analysis could then be conducted to calculate confidence intervals for key decision criteria such as total GVA and expected outputs. Yet, such (decision) modelling is ultimately dependent upon sound evidence and evidence synthesis and is still couched within the political nature of using such decision tools.

It will be interesting to see if the WWG can meet its ambitious mandate given the theoretical, methodological and practical difficulties highlighted above. Regardless, it could be argued that the WWG does not go as far as NICE in terms of generating a pull for evidence centrally, formally peer-reviewing such evidence with identified stakeholder groups and actually incorporating such evidence directly into decision making via an institutional process (with political backing and a legislative framework).

6.6 Conclusion

This chapter has contributed in a number of ways to address the first research question. It has demonstrated that with the formation of the Coalition government there was seemingly a wave of enthusiasm for experimental study designs and extending approaches more closely aligned with EBM to wider social policy evaluation. By mirroring analysis of the generation and use of evaluation evidence across the

⁸⁴ More frequently employed within EBM, see Higgins, Thompson and Spiegelhalter (2011).

health and regional policy contexts it has been possible to highlight their differences to reveal how parallel EBM processes for evaluation and investment prioritisation are played out in different ways within each context. It was found that there are a number of implicit assumptions embedded separately within EBM and within EBRPM which have direct implications for an extension of an EBM approach to regional policy (and to wider health and social policy) and significant epistemological, methodological and practical difficulties were identified.

Reflecting upon this, it is perhaps unsurprising that it was also found that claims that RCTs and quasi-experimental approaches can and should be applied within spatial policy (BIS 2014b) are already beginning to unravel in reality as the evaluations of new local growth initiatives are being commissioned. Likewise, the WWG has faced significant challenges in applying systematic review and meta-analysis to the current spatial policy evidence base. It could be concluded that, in some ways, there has been an opportunity missed in extending the EBM approach. A more nuanced review of the vast literature on EBM, the methodologies employed and a greater understanding of the NICE process, alongside a greater appreciation of the political nature of decision making, could have provided a richer insight into the appropriate use of evidence within regional policy making. This stands in contrast to simply appealing to the generation of particular forms of evidence. This insight provides a foundation for the following chapters to build upon.

To conclude this comparative literature review, Chapters 4-6 have drawn upon academic and policy literature to explore the implications of applying an EBM approach to regional policy evaluation. Strong emphasis has been placed on reviewing debates focused on the generation of evidence and the credibility of certain evidence types through reviewing the methodological guidelines for economic evaluation and the central 'pull' for evaluation evidence for investment prioritisation. Therefore, for the next empirical part of this thesis, the focus will be on how evidence was incorporated into regional policy organisations (the RDAs) and the role of other factors besides evidence. Such investigations uncover contextual factors which influence the process of evaluation generation, communication and use within regional policy making. In order to frame this analysis, the next chapter provides an overview of the policy making processes of the RDAs.

Chapter 7

The RDA Organisations: An Example

7.1 Introduction

In this chapter the evaluation and policy making processes of an example RDA, Yorkshire Forward (the RDA for Yorkshire and the Humber), are described. The purpose of this chapter is to give background, contextualising information to position the findings of the online survey and online workshop, which are presented in the following empirical chapters.

7.2 The principles of RDA policy making

The Labour Party Manifesto (1997) set out a mandate for the RDAs to “co-ordinate regional economic development” and an overarching vision for Evidence Based Regional Policy Making (EBRPM) can be inferred from the principles of “accountability, effectiveness and subsidiarity” set out in the 1996 ‘Report of the Regional Policy Commission’ (cited by Hayward 1997, p378). Firstly, in terms of an accountable policy making process, a paper by Blagescu, de Las Casas and Lloyd (2005, p4) is of interest. They note that accountable organisations are: transparent; engage the participation of stakeholders; evaluate performance and disseminate evaluation outputs; and provide a feedback mechanism for stakeholders. In terms of an effective policy making process, ‘Professional Policy Making for the Twenty First Century’ (Cabinet Office 1999b, p13-14) identified the need for: systematic evaluation for effective policy making; to use the best available evidence from a wide range of sources; and to learn from experience of what works and what doesn’t. Armstrong and Taylor (2000) agree that evaluation is essential for regional policy to be efficient, effective, and to meet its objectives (including economic, social and environmental objectives).

7.2.1 The assessment of RDA success

Regionally, the RDAs were initially accountable to indirectly elected Regional Chambers made up of regional partners, including “representatives of local authorities, economic and social partners (e.g. business associations, trade unions and voluntary groups) and other sectoral interests (e.g. higher education, environment and rural)” (Pike et al.

2012, p104). RDAs were also directly accountable for the way in which they used their resources nationally through their sponsor department (BIS) and for delivering effectively against Regional Economic Performance (REP) Public Service Agreement (PSA) monitoring targets set by central government. In particular, PSA7 tasked the RDAs with improving the economic performance of all English regions and reducing the gap in economic growth rates between regions (HM Treasury 2003a). The most recent reporting regime monitored performance in terms of ONS Regional Accounts GVA estimates (ONS 2011).

It was found in Chapter 5 that, prior to the RDA national impact evaluation, there was a complex model of RDA accountability and that it lacked an identified point in the policy process when the integration of the multiple forms of evaluation evidence being produced by the RDAs would be expected to be reported back to central government. At an operational level, delivery of PSA7 was implemented via Regional Economic Strategies (RES). RDA's were also considered an appropriate institutional framework to operate 'indigenous development' policies and thus their objectives were wide-ranging: "furthering economic development and regeneration; promoting business efficiency and competitiveness; promoting employment; enhancing the development and application of skills relevant to employment; and contributing to sustainable development" (Great Britain 1998, p8). Often the RDAs sought to achieve their objectives via funding projects through local level 'delivery' organisations, as a means of enabling the active participation of the local community. Polverari and Bachtler (2004, p12) note that the number of actors and mechanisms involved in policy making became "unprecedented in comparison with the past."

In 2004 plans to enhance the accountability of regional institutions, through an elected regional assembly, were rejected and subsequently the 'Sub-National Review of Economic Development and Regeneration' (SNR) process was introduced in 2007 to streamline state involvement in regional policy (Pike et al. 2012). The sub-regional level was already prominent in Yorkshire and Humber at this time, with sub-regional investment plans integrating Structural Funds, RDA and Local Authority funding streams for economic development and skills. The SNR process placed more emphasis upon sub- and city-regional partnerships and joint working, Regional Ministers, a parliamentary regional select committee (Pike et al. 2012), the delivery of PSA7

through integrated strategies (HM Treasury 2007), and led to the abolition of the Regional Chambers. Leading up to the UK General Election, Pike et al. (2012, p104) notes that “SNR collided with the assessment and emergent critique of New Labour’s approach.” The first RDA national evaluation was published (PWC 2009a) and, despite presenting broadly positive conclusions on the effectiveness and efficiency of the RDAs, Ferry and Bachtler (2013, p269) note that “these were ignored and indeed contradicted by the political narrative.”

7.3 The case study of Yorkshire Forward (YF)

This section is primarily based upon two papers that were produced as part of the ‘Learning Legacy’ series during YF’s transition to closure: ‘Research, Intelligence and Evaluation’⁸⁵ (YF 2011c) and ‘Economic Strategy’ (YF 2011b). Although not academic, peer reviewed papers, these give an ‘insider’s account’ of RDA evaluation and the inner workings of RDA policy processes. In addition, the YF Regional Economic Strategies (RES), Corporate Plans and Annual Reports were reviewed.

7.4 Yorkshire Forward: an example of EBPM?

YF’s evolving approach to policy making, and the role of evidence within policy making processes, can be traced back by reviewing the development of the Regional Economic Strategies (RES) for Yorkshire and the Humber over time. The ‘Research, Intelligence and Evaluation’ legacy paper highlights that for the first RES (2000-2006) an evidence based approach was constrained by the “paucity of available intelligence” (citing the 1999 ‘state of the region’ report, YF 2011c, p3). However, by the time of the second RES (2003-2012), “the executive summary... proclaimed it to be an evidence based strategy” (YF 2011c, p3). Demonstrating further commitment to an EBPM approach, at least at a rhetorical level, the final RES (2006-2015) set out an aspiration for YF to “be at the forefront of intelligence and evaluation activity in the United Kingdom” (YF 2011c, p3).

⁸⁵ Drawing upon desk based research, the legacy paper goes on to describe the development of a range of discrete but interlinked activities undertaken by the YF’s research, intelligence and evaluation functions. Detail is given on (YF 2011c, p4-14): Yorkshire Forward’s Chief Economist Unit (CEU); The Yorkshire Futures regional observatory; City and Sub-Regional Analysts; Office for National Statistics (ONS) regional presence staff; and Yorkshire Forward’s Evaluation Team.

The rationale for YF's investment in the development of regional intelligence⁸⁶ is set out in the 'Research, Intelligence and Evaluation' legacy paper as having aimed to: identify the key issues faced by the region; design projects and programmes to meet identified needs, based on what is known to work; prioritise resources; learn from experience; improve delivery of interventions; understand and evidence the impact of interventions; be accountable to regional, national and local stakeholders; monitor change in the region; and anticipate future change and prepare accordingly (YF 2011c, p3). This implies an aspiration for regional policy making which is evidence based (or at least evidence informed), accountable and effective. However tension is revealed as the paper 'Economic Strategy' highlights the importance of other factors besides evidence such as politics and regional partner's values and preferences (YF 2011b, p2):

Good strategy = evidence [&] preference. The best strategies and plans do things that are wanted by partners and backed by evidence. The more this is achieved the greater the chance of lasting impact, even if the real world balances will inevitably have to be struck.

The tension between central and local relations, as explored by others (Benneworth 2011), was highlighted in the same report (YF 2011b, p16):

BIS... push[ed] for rigid channelling of RDA resources into activities that yielded the best GVA and value for money returns. In contrast local partners wanted to do what they felt was right for their area based more on judgement and strategic added value.

7.5 Policy making processes in Yorkshire Forward

7.5.1 Organisational structure

YF was run by an executive board, and had both a chair and a chief executive. YF's chair and chief executive were senior figures who represented the region at national and parliamentary levels (YF 2009). The board was made up of senior leaders from local government, the voluntary sector, trades unions and local businesses (YF 2009). The aim was for each member to 'champion' areas of YF policy, and to maintain strong relationships with regional partners and stakeholders (YF 2009). The board of directors met every 6 weeks and had responsibility for the overall strategic direction and

⁸⁶ Defined to be: "primary and secondary research; monitoring; evaluation; modelling; policy analysis; and analysis of private and public sector data sets" (YF 2011c, p3).

management of the RDA (YF 2009). In the YF Annual Report 2010-2011 (YF 2011a, p9) it was noted that the Chief Executive presented a 'Progress Report' at every Board meeting, accompanied by the latest 'KPI [Key Performance Indicator]', 'Significant Risks' and 'Significant Issues' reports. The Chief Economist's Unit (CEU) had a role in producing economic briefings for the Board (YF 2011c, p4) and the economic downturn increased the demand for these (YF 2011c, p5). It could be inferred that this reporting framework was essentially managerialist rather than focusing on strategy. Another weakness was that the Board members did not necessarily know about 'what works' beyond their own experience.

YF had a senior management team and was organised into 5 delivery directorates, each led by an executive director including: Business, Economic Inclusion, Environment, Finance and Strategy alongside a corporate management team. The Strategy Directorate is of particular relevance and included the following teams: CEU Economic Policy & Strategy (including evaluation); Sustainable Development; Transport; Yorkshire Futures. This structure was quite similar to other RDAs.

7.5.2 Strategy

The RDAs were financed by national government public funds via the creation of a 'single pot' of RDA funding. There was a high degree of flexibility, enabling funding to be targeted towards the investments the RDAs prioritised for each region (YF 2009). At their establishment, RDAs were tasked to formulate and keep under review a Regional Economic Strategy (RES) to guide delivery on business, employment, skills, regeneration and sustainable development (Great Britain 1998). YF produced three RES's, and although the process evolved over time, it generally involved: drawing upon an evidence base and baseline; developing a strategic vision; undertaking rounds of consultation with regional stakeholders and refining the RES; peer review/sustainability appraisal; endorsement by RDA Executive/ Board and central government (YF 2011b, p5-18). Every 3 years, YF had to agree a corporate plan with BIS which set out objectives showing how YF would deliver their RES (YF 2011b). The Yorkshire and the Humber Regional Committee stated that the last RES (2006-2015) "was drawn up after a comprehensive and iterative process of discussion and negotiation with relevant stakeholders, but on a top down basis" (2010, p71). The 'Sub-National Review of Economic Development and Regeneration' (SNR) process

identified that the upcoming 'Integrated Regional Strategy' in 2009 should be "built from local priorities" and "built from the bottom up,"⁸⁷ placing emphasis on closer joint working with regional stakeholders and partners.

7.5.3 Project appraisal

YF's 'Performance Management Framework' (PMF) provided a process for the project cycle including development, appraisal, monitoring and evaluation. The PMF aligned with Guidance for RDAs in Appraisal, Delivery and Evaluation (GRADE), issued by BIS. Principles for the PMF process were set out as ensuring that YF: delivered value for money for the public purse; complied with Government requirements; and learned lessons (YF 2011a, p22). In practice, the 'Economic Strategy' legacy paper (YF 2011b, p16) notes that project ideas "identified as desirable under any of the action planning type systems over the years" then had to be worked up as full proposals and managed to completion under YF's PMF. There were changes and improvements to the system over the years along with substantial learning. The Strategy Team led a strategic appraisal function and chaired weekly strategic appraisal panels which also brought in other directorates and for some years external partners at Government Office and the Regional Assembly. "Robust appraisal, judgement, pragmatism and communication" were identified to be important (YF 2011b, p16). In addition the CEU undertook economic appraisal work for major projects "to help shape and make the case for major projects undergoing government appraisal" (YF 2011c, p5).

7.5.4 Project delivery and monitoring

The Finance Directorate then took the lead on subsequent (and more detailed) Full Business Plans and matters of project delivery and monitoring. The monitoring process included: initial review meetings; quarterly reports submitted to the RDA from the contractor followed by quarterly review meetings; verification visits; and annual reviews. Project managers also often had more regular contact with contractors that would include phone calls, email correspondence and face to face meetings when required.

⁸⁷ Local Government Yorkshire and Humber Sub National Review Officers Group (2009).

The YF 'Artemis' IT database was used to support the PMF process and to collate financial and output monitoring data on projects and programmes. Financial and monitoring data collated was based upon the RDA tasking framework (cited in DTI 2006, p141-143). Core outputs included: jobs created or safeguarded; people assisted into employment; skills assists; businesses created; businesses assisted; and land remediated. Although Artemis output monitoring data was often analysed when conducting evaluations of YF's investments, it was primarily used for financial reporting and the monitoring of spend against budgets by directorate and overall. Indeed, reflecting upon the utilisation of the system, the NAO commented (2007) that "Yorkshire Forward's systems for monitoring expenditure have been successful in ensuring a more predictable and even spend profile over the financial year." Unfortunately, YF's legacy paper series did not comment on the lessons learnt from the ICT monitoring systems at YF. However there were recorded issues. For instance, problems arose during 2004-05 in monitoring the split of expenditure between current and capital elements for ERDF (NAO 2007).

The 'Research, Intelligence and Evaluation' legacy paper (YF 2011c) made the point that the design of monitoring and management information systems should meet the needs of evaluators as well as project managers. The need for a more holistic approach to monitoring change was noted to enable monitoring of social, environmental and economic factors as well as integrating management information systems for recording details of beneficiaries of interventions, with intelligence functions such as official statistics, commercial data sets and evaluation. It was noted that for YF, work with a GIS system (Geographic Information System) had gone some way towards this but that a 'Regional Knowledge System' (a system recording business support interventions) was in the early stages and had yet to be fully exploited.

7.6 The generation of evaluation evidence in Yorkshire Forward

As described in the 'Research, Intelligence and Evaluation' legacy paper (YF 2011c, p14-17), in 2004 YF was "one of the earlier RDAs to increase its evaluation capacity and investment." An evaluation manager was recruited who then developed the first YF evaluation strategy. This placed emphasis on the role of learning, project-level evaluation and internal evaluation. In 2006 a new evaluation strategy was produced and in 2009 the strategy was amended. These took into account publication of the IEF

(DTI 2006) and IEF + (BIS 2009) respectively. The legacy paper goes on to document four “major challenges” for YF to meet the requirements of the RDA national impact evaluation including (YF 2011c, p16): the difficulty of meeting the coverage target of 60% of expenditure to date when evaluations were planned at a project level; the underestimation of RDA impact due to value placed on net jobs but not on skills or social development; issues of aggregation due to inconsistent data across and within the RDAs; and the lack of value placed on internal evaluation and learning. The first three points have previously been reflected upon in Chapter 5.

Another factor which influenced evaluation strategy and planning was that from 2007 onwards, delivery at YF began to move towards a programme approach from a project level approach.⁸⁸ However by 2010, “programme level working had not been fully embedded, so evaluations remained a complex mix of project, programme and thematic studies” (YF 2011c, p17). Despite the challenges, the legacy paper (YF 2011c, p17) states that for evaluation “by 2010, major advances had been made in implementing a high quality and coherent strategy.” Overall, a commitment to the production of evaluation evidence with a “strong focus on capturing lessons learned from investments” was set out in the YF Annual Report 2010/11 (YF 2011a, p3).

In terms of evaluation practice, activities conducted by YF’s evaluation team focused on the generation and dissemination of evaluation evidence, were wide ranging and included (YF 2011c, p15): developing and leading on the implementation of an evaluation strategy for YF (single pot and ERDF funded activities); developing evaluation plans for YF funded projects/programmes with project managers and partner organisations; reviewing project proposals (for adequacy of evaluation plans and application of learning); commissioning and managing independent evaluations; procurement and management of a panel of external evaluators; conducting internal evaluations; training and awareness raising activities; and disseminating evaluation findings through published reports, summaries, events and briefing sessions.

The Evaluation Team were expected to provide a central source of expertise to the RDA and to ensure high quality standards (YF 2011c). The ‘Research, Intelligence and Evaluation’ legacy paper describes the importance ascribed to the evaluation team

⁸⁸ Policy product ranges and geographic programmes were clearly specified.

being independent from project delivery functions for accountability purposes, particularly for commissioning external evaluations (YF 2011c, p15). For external evaluations of projects/programmes, it was expected that a (broad) range of evaluation evidence would be produced including (YF 2011a, p3):

- Whether objectives are being met;
- The net economic impact of interventions;
- The social and environmental impacts of interventions;
- The value for money, including return on investment of interventions;
- The strategic added value of interventions;
- The effectiveness and efficiency of delivery of interventions;
- Lessons learned and good practice to share from interventions.

The key point to be made here, however, is that evaluations were generally project level and followed a set script.

7.7 The communication and uptake of YF evaluation evidence

In terms of mechanisms for internal communication and cross- directorate working within YF, the YF Annual Report (YF 2011a) identified a number of ‘channels’, including: team meetings; core team briefings for the communication of strategic messages; ‘all staff’ emails; a staff magazine (‘Ontrack’); a staff intranet (YFi); field trips; staff annual survey; informal questioning sessions for Executive Directors’ (‘surgeries’); informal questioning and presentation sessions for the Chief Executive (‘Roadshows’); internal communications forum; and a Staff Conference. The ‘Research, Intelligence and Evaluation’ legacy paper (YF 2011c) did not comment specifically on the use of these mechanisms to communicate evaluation findings internally.

For external dissemination, the legacy paper (YF 2011c) discussed the use of ‘the What Works database.’ This was essentially a knowledge management repository: a searchable library of case studies, evaluations and research developed by Yorkshire Futures. Although not discussed in the legacy paper (YF 2011c), the work of the Office of Project and Programme Advice and Training (OffPAT) was along a similar vein. The OffPAT e-library provided a shared repository where RDAs could place completed evaluation reports and OffPAT would then produce a short executive summary of each evaluation.

Overall, the legacy paper (2011c) called for a more active dissemination of RDA evaluation findings. Three key ways in which ‘generic lessons’ could be improved were identified as follows (YF 2011c, p23): “increasing awareness of evaluation findings among personnel with a cross-organisational perspective (such as appraisal staff and contract monitoring staff); developing libraries and repositories of resources such as the Yorkshire Futures ‘what works’ database”; and “improving the synthesis of evaluation lessons.”

In terms of the uptake and use of evidence, the legacy paper notes positive feedback on the approach to self-evaluation and the development of an ‘enabling’ evaluation culture (GHK 2008 cited by YF 2011c, p20). Nonetheless, the limited use of evidence in decision making is clearly articulated: “many of those engaged in research and intelligence expressed a view that the region was still some way from ‘evidence based’ policy making. It was rare to see examples of where an intervention was shaped principally by evidence. Nevertheless, evidence undoubtedly had an important role to play” (YF 2011c, p20). Yet what the role of evidence was within policy making processes, and the magnitude of that role, were questions that remained unanswered within the RDAs (YF 2011c, p3):

The relationship between intelligence and resulting actions is a complex one that is rarely articulated, documented or evaluated. Hence, it is difficult to assess objectively how successful an investment in intelligence resources has been.

7.8 Conclusion

This chapter has described the evaluation and policy making processes of an example RDA, Yorkshire Forward, to give background, contextualising information to position the findings of the online survey and online workshop (as discussed in the following chapters). Drawing primarily upon two RDA legacy papers, which gave an ‘insider’s account’ of RDA evaluation and the inner workings of RDA policy processes, this chapter has demonstrated that the uptake and use of evidence within RDA policy making processes was not widely understood. In addition, a need was identified for the active dissemination of RDA evaluation findings. This provides a foundation for the following chapter that will explore the way in which evidence was incorporated into

the policy making processes of the RDAs, including the communication and use of evaluation evidence, as well as the role of other factors besides evidence.

Chapter 8

The RDA Evaluation Experience: A Case of Evidence Based Regional Policy Making?

8.1 Introduction

This chapter addresses the second research question: what factors influenced the generation, communication and use of evaluation evidence within the English RDAs? This is explored through analysis of the perspectives of an expert stakeholder group and builds upon the groundwork provided in Chapters 4-7.

An online survey was conducted with an expert stakeholder group to capture the perspectives of personnel engaged in RDA policy evaluation. Responses were elicited from commissioners, producers, and users of evaluation evidence across the policy cycle. The theoretical background for the chapter is initially reflected upon, drawing upon the literature review. Then the characteristics of the survey respondents are discussed. The rest of the chapter is then structured by the themes that emerged from the literature review (Chapter 2), from analysing the policy making processes of the example RDA (Chapter 7) and the survey data including: the uptake and use of evaluation evidence within RDA policy making process; the generation and communication of RDA evaluation evidence; and the factors that influenced the generation, communication and use of evaluation evidence. Finally, the potential role for knowledge translation tools to increase the utilisation of RDA evaluation will be considered.

8.2 Theoretical background

This part of the empirical research predominantly focuses on debates around the way in which evidence was incorporated into the policy making processes of the RDAs, including the communication and use of evaluation evidence, as well as the role of other factors besides evidence. Therefore the unabridged discussion of the theoretical background for this chapter is provided in Chapter 2 (predominantly sections 2.4 and 2.5).

It was found in Chapters 4-6 that there was evidence of RDA regional policy evaluation guidance shaping an implicit evidence hierarchy whereby macro, quantitative approaches (such as experimental surveys) are placed at the top, and micro, more qualitative approaches (such as case studies and partner consultation) are placed at the bottom. An inherent assumption can be deduced that study design and internal validity were the key indicators denoting the strength of evidence. The institutional framework for incorporating evaluation evidence into policy making and the division between central and regional agencies was also found to be somewhat unclear. The RDA national evaluation was the first attempt to provide an independent assessment of the net impact and economic value of RDA interventions. More generally, the regional policy context was found to be highly complex given its multifaceted policy agenda, structures and mechanisms, alongside its inherently political character.

Reflecting upon the above findings and drawing upon the EBPM/KTE and political science literature review (Chapter 2), it could be hypothesised that the RDA evaluation guidelines may have led to a central 'pull' for evaluation evidence focused on certain types of knowledge (i.e. that which is "derived through quantitative methodologies, empirically-tested and validated") which is seen to lead to "instrumental rationality" and a managerialist and mechanistic approach to policy making (Sanderson 2002, p6). Similarly, Rutter, Hawkins and Parkhurst (2013) discuss the concept of 'Evidence Controlled, Managed and Legitimised Policy' (ECMLP) rather than evidence based policy. When considering the fundamentally political nature of evidence, the question for policy makers is often, therefore, not simply 'what works', but "what is appropriate in the circumstances, and given the overall policy objectives" (Rutter, Hawkins and Parkhurst 2013, p 17). As discussed in Chapter 2, Huber's work (2006) is apposite here as he described evaluation being used as: 'window dressing', a 'formal exercise' or part of a 'co-ordinated learning process.'

It was also found that it is often unclear how evidence is incorporated and used within the policy making processes of organisations (Chapter 2). Indeed, less attention has been paid to the evaluation of UK regional policy instruments (i.e. non EU policy instruments) and to the processes of undertaking regional policy evaluation in practice (Chapter 1). This chapter seeks to explore the perspectives of an expert stakeholder

group to understand the application of the regional policy evaluation guidelines and the central 'pull' for evaluation evidence within the RDAs.

8.3 Who were the survey respondents?

Methodological considerations, including discussion of empirical data collection, analysis and interpretation are presented in Chapter 3 (section 3.6). In particular, see section 3.6.1 for discussion of the survey respondents.

8.4 Presentation of the findings

Given that the RDA sample was biased towards Yorkshire Forward (YF) personnel in particular, analysis was undertaken to compare the responses from YF and 'other RDA' organisational subgroups. As analysis of the quantitative data did not suggest heterogeneity between sub-groups, the quantitative results are presented for the total population when reporting the survey findings in the narrative. For information purposes, the quantitative results are also presented for the YF sub-group through the use of footnotes so that the thread of the narrative is not compromised. For the qualitative data, quotations are presented denoting the departmental sub-group and identification number of the respondent. For reasons of anonymity (given the small sample sizes of departmental sub-groups), and given that analysis of the quantitative data did not suggest heterogeneity between sub-groups, the qualitative data across organisations has been merged and does not separate YF personnel from other RDA personnel.

For further information: survey items and respondents are presented in Appendix 2; the quantitative data are presented in Appendix 5; the qualitative data are presented in Appendix 6.

8.5 The principles of regional policy making and the role of RDA evaluation

The overarching principles for regional policy making were set out in section 6.2 and the role of RDA evaluation in ensuring accountability and effectiveness in RDA policy making were themes present within the survey data. In terms of accountability, a regulatory responsibility for the RDAs to comply with BIS and EU directives was underlined in the data. Respondents also identified that the funding of the RDAs via

the 'single pot' and ERDF (i.e. by public money) meant that the RDAs had both a financial, and a moral, responsibility to wider stakeholders:

There is an obligation... for [the] public sector to be able to account for decisions made with tax payer's money. (External evaluator [75])

The survey revealed the perception that evaluation had played an important role in supporting the RDAs to be 'accountable organisations.' Most survey respondents (84%, 68)⁸⁹ agreed that undertaking RDA evaluation had 'showcased the effectiveness of RDA interventions to internal and external audiences.' Nevertheless, an underlying vested interest to demonstrate organisational success was identified:

Government got into the habit of giving, then taking, responsibilities... [the RDA was] a funding agency and not really a part of the core mission. (External evaluator [76])

[Some] RDAs really didn't want objective evaluations, just good news. (External evaluator [71])

Most respondents (84%, 67)⁹⁰ agreed, however, that RDA evaluation was conducted in a way that 'ensured the independence of evaluation' outputs and the role of independent (i.e. external) evaluation evidence was emphasised:

[Evaluation] provides compelling independent evidence of what does and does not work... (RDA economic appraisal officer [5])

This concept of 'what works' was a central theme present within the qualitative data. In particular, respondents identified the potential for evaluation to influence the effectiveness of RDA policy making processes. Normative statements by respondents emphasised the importance of evaluation for organisational learning and innovation, and this was often given as the principal rationale for investing resources into RDA evaluation:

Evaluations were important for learning and continuous improvement. (RDA evaluation officer [51])

Public money should be spent on evaluation to improve future performance and impact. (RDA strategy officer [64])

⁸⁹ YF: 88%

⁹⁰ YF: 88%

In contrast, the quantitative survey data revealed that, despite most respondents (85%, 69)⁹¹ agreeing that RDA evaluation ‘highlighted what works and what does not work,’ only half (54%, 44)⁹² agreed that evaluation ‘enhanced the effectiveness of RDA performance and effectiveness.’ Of particular significance is that only a third of respondents (35%, 28)⁹³ agreed that RDA evaluation processes ‘ensured learning and development from evaluation findings.’ This will be investigated further later in the discussion.

Overall, a key finding of the survey is that RDA evaluation was perceived to have a greater influence on demonstrating accountability than on enhancing the effectiveness of RDA policy making.

8.6 The uptake and use of evaluation evidence within the RDAs

The policy making processes of the RDAs involved a range of discrete but interlinked activities including strategic decision making, project appraisal and delivery (see section 7.5). Yet the influence of evidence on such policy making functions was not widely understood within the RDAs (YF 2011c, p3):

The relationship between intelligence and resulting actions is a complex one that is rarely articulated, documented or evaluated. Hence, it is difficult to assess objectively how successful an investment in intelligence resources has been.

The survey explored the relationship between evidence and policy action further.

8.6.1 Strategic decision making

Central financing of the RDAs, including the prima facie flexibility of ‘single pot’ funding, was discussed previously, and a key area that the survey aimed to investigate further was the role of evaluation evidence within RDA investment prioritisation activities at the strategic level. Responses from the survey suggested that investment prioritisation between broad policy areas (i.e. allocating investment between programmes) was undertaken by Executive teams (i.e. senior management) within the RDAs. A significant challenge for prioritising investment to maximise impact was

⁹¹ YF: 85%

⁹² YF: 60%

⁹³ YF: 29%

deduced in that the diverse nature of projects and programmes run by the organisation meant that comparison between them was difficult (see section 9.5.1 for further discussion).

In addition, the dominant pressure within the RDAs to meet spend targets was identified:

Often spend was the main focus... Getting money out the door, especially if close to year end... (RDA project delivery/performance officer [12])

Although 64% (52)⁹⁴ of survey respondents agreed that evaluation 'provided evidence for RDA investment prioritisation,' there was a sentiment in the qualitative data that there had been a missed opportunity for evaluation evidence to be fed through more systematically into this stage of the policy process:

We didn't really get evaluations to feed into internal and external strategy development work in terms of what works best and cost effectively at delivering outcomes. (RDA strategy officer [64])

[There was] haphazard investment. (RDA project delivery/performance officer [29])

That economic appraisal and evaluation evidence played a minor role in strategic investment prioritisation compared to other factors such as political strategies, previous practice, ideas and values, was a prevalent theme:

[Investment prioritisation was] ad hoc, driven by Government, local political pressure, internal personalities and response to economic events. (RDA strategy officer [60])

There was some evidence that evaluation evidence was used to justify decisions that had already been taken:

It was unclear whether evaluation work was being used to justify projects or guide project development. (RDA economic appraisal officer [10])

There are exceptions to this but normally CBA [cost benefit analysis] occurred when a project had to go to central Government and then the agency took it seriously (though it was still an exercise in justifying something that others had already decided should happen). (RDA strategy officer [63])

⁹⁴ YF: 63%

Evaluation evidence was presented but I am not sure how much it actually influenced the decisions made. Evaluation evidence was perhaps just used to justify the decisions made. (RDA evaluation officer [53])

Much of the evaluation completed to 2010 had been skewed towards justifying the RDAs. (External evaluator [72])

It was identified that sometimes evidence was therefore ‘cherry-picked’ to support political decisions:

If anything the use of evaluation by the senior team was more an exercise in politics (using evidence to talk up the RDA or seeking to undermine it where it provided a more negative conclusion). (RDA strategy officer [63])

This finding aligns with Huber’s (2006) terminology of evaluation being used as ‘window dressing,’ giving a veneer of credibility to policy processes externally. When reflecting upon the lack of evidence informing strategic investment prioritisation, statements by respondents were generally negative in tone:

The maintenance of politically driven investment schemes with little evidence probably doomed the RDAs. (Central Government officer [82])

The influence of evaluation evidence on RDA strategic decision making (SDM) more generally was also explored within the survey and 74% (60)⁹⁵ of all respondents agreed that evaluation evidence ‘clarified objectives and strategies for decision making.’ An additional quantitative survey item, targeted towards personnel involved with RDA strategy processes,⁹⁶ aimed to uncover the influence of evaluation evidence on other potential strategic ‘uses’ of evaluation evidence (Appendix 2, question 20). Approximately half of these respondents (53%, 20) agreed that ‘evaluation evidence fed into the RDA’s Corporate Plan.’ Yet, when regarding more routine SDM processes, only a third of respondents (37%, 14) agreed that ‘evaluation evidence fed into decision making at Executive meetings’ and a quarter (24%, 9) agreed that ‘evaluation evidence fed into decision making at Board meetings’. The complexity of incorporating

⁹⁵ YF: 77%

⁹⁶ The influence of evaluation evidence specifically on strategic decision making was elicited through targeted questions for respondents involved with RDA strategy, including strategy officers (n=8), evaluation officers (n=14), external evaluators (n=13) and central government officers (n=3) (Appendix 2).

evidence into SDM, alongside other factors such as politics and regional partners' values and preferences, was identified:

Were the RDAs there to make strategic decisions or implement policy that had some democratic accountability regardless of what the evidence says? (External evaluator [74])

Still a more common theme was the suggestion that evaluation evidence 'should' have played a greater role in SDM processes, with ideas put forward by respondents for 'getting evidence into policy' emphasising the need for a more integrated approach and broader application of evaluation evidence:

Evaluation evidence should have informed decision making and strategic planning at all levels within the RDA, for example: Regional Economic Strategy, investment allocation between... Programmes... [and] partnership decisions. (RDA strategy officer [59])

In turn, the types of evaluation evidence that would have been relevant for SDM were reflected upon, and RDA evaluators identified the importance of longitudinal, programme and thematic evaluations to draw out key lessons and Strategic Added Value (SAV). RDA evaluation in its existing form was found to have influenced SDM to an extent, but in less direct ways. This supports work done by Nutley, Powell and Davies (2013) who found that evidence may be used in subtle ways within policy processes. For instance, informal mechanisms to share learning from evaluation for SDM were described by one respondent:

Evaluation evidence often informed strategies and plans in an informal way. For example, senior managers would have been made aware of significant evaluation findings and would then be aware of these findings and their implications when drafting strategic documents. (RDA strategy officer [59])

8.6.2 Project appraisal

Activities undertaken for the strategic and economic approval of RDA projects were discussed in section 7.5, and a key area that the survey aimed to explore was the role of evaluation impact data and qualitative lessons influencing RDA project appraisal. Overall, it was found that the subordinate role of evaluation evidence within RDA strategic investment prioritisation, discussed above, was then reflected within RDA project appraisal processes.

In terms of economic project appraisal, the majority of all respondents (90%, 73)⁹⁷ agreed that evaluation ‘provided evidence for the economic appraisal of individual projects.’ However, the emphasis in the survey item was on supplying, rather than on utilising evidence. It was emphasised by RDA economic appraisal officers that usually economic appraisal work would be focused towards major projects undergoing government appraisal, rather than for all RDA investments. Given that most RDAs took a project (rather than a programme) approach to delivery, and given that the RDAs often prioritised investment on an ad hoc basis rather than through formal funding rounds, projects and programmes were seldom directly compared against each other. Instead a pragmatic position towards economic appraisal was identified in the qualitative data:

There would never be a way of consistently appraising or evaluating every project. (RDA economic appraisal officer [10])

This was perceived to be a missed opportunity:

[Economic appraisal] should have been undertaken as part of a genuine business case for investment and not as a tool to justify a decision already made. (RDA strategy officer [63])

In turn, the types of evaluation evidence that would have been relevant for economic appraisal decision making were reflected upon, and frequently the potential for using benchmark data was identified by RDA and central government officers:

Benchmarks... would have been useful in appraisal, and could have been used as a guideline to strive for higher Value for Money. (RDA project delivery/performance officer [15])

Individual RDAs had a strong evidence based approach and used a basket of ready reckoners and associated models to demonstrate highest impact. (Central Government officer [82])

Although some RDAs systematically used quantitative GVA benchmarks (BIS 2009c) for economic appraisal, some respondents cautioned that such an approach was constrained by data quality and relevance issues. There was a lack of consensus over whether the use of benchmarks was an ‘appropriate’ use of evidence:

⁹⁷ YF: 92%

Too many academic arguments about who was right and who was wrong about the use of [benchmarks] and a reluctance to start using them because they were a "blunt tool." However there were never really any alternatives put forward. [The RDA] took a position whereby we wouldn't use them as a result - other RDAs took the opposite view. (RDA economic appraisal officer [5])

In terms of strategic project appraisal, it was anticipated that the survey might identify that evaluation evidence was applied to shape project design, with a focus being not only on 'what works' but on why certain interventions work in certain circumstances and for certain groups of people (linking to work by Pawson and Tilley 1997). Although 60% (49)⁹⁸ of all respondents agreed that evaluation evidence 'improved project design and development,' the qualitative data suggested that evaluation evidence was often simply quoted to meet the requirements of the funding approval process:

Evaluation evidence was often cited in investment appraisals more as a placeholder than as a serious consideration for projects. (RDA evaluation officer [53])

Tacit knowledge produced by evaluation on best practice and lessons learnt being primarily used for 'exemplification' in the RDAs, rather than for shaping project design and future delivery, was also found by Cook et al. (2008). The principal reason cited for this was the lack of policy relevance (i.e. generalisability) of evaluation evidence to the appraisal of future investments:

The evaluation team's input to the appraisal process involved highlighting and applying evaluation findings of previous projects to current projects. This was often difficult, however, as evaluation findings were often quite specific and rarely fully applicable to other projects. (RDA economic appraisal officer [3])

The cursory use of evaluation evidence at this stage of the policy process aligns with Huber's (2006) terminology of evaluation being limited to the minimum required for a purely 'formal exercise.' Supporting this, content analysis of the qualitative data revealed that incorporation of evaluation evidence into investment appraisal was perceived to be "bolted on" (n=3), seen as an "add on" (n=8) or part of a "tick box" exercise (n=8) rather than an integral part of the policy process.

Instead, the dominant influence of other factors besides evaluation evidence was once again identified:

⁹⁸ YF: 60%

The organisations' spending processes did not require the same scrutiny as national projects and I think decision makers were more influenced by political factors in project choice than value for money indicators. (RDA economic appraisal officer [10])

In turn, ideas for increasing the awareness and application of evaluation findings for project appraisal were reflected upon by respondents:

Evaluations should have fed into a framework/matrix which could have been used to influence the appraisal process to a greater extent. (RDA project delivery/performance officer [28])

8.6.3 Delivery

The activities undertaken for project/programme delivery in the RDAs were wide ranging (section 7.5), but it is the activities of monitoring and evaluation that are the focus here. Although less than half of all respondents (46%, 37)⁹⁹ agreed that RDA evaluation 'increased management and delivery efficiency,' finer distinction between respondents is needed. The qualitative data identified that the extent to which evaluation evidence was perceived to influence delivery was directly linked to the level of engagement that delivery partners and RDA project managers had with the evaluation process:

On the whole, project managers were very supportive once they realised what was involved and how valuable the findings were. (RDA evaluation officer [58])

The quality of the relationship (i.e. building trust and collaboration) between evaluators and evaluation users was identified to be important within the research process, and RDA officers highlighted the time taken to build and maintain such relationships. This aligns with findings from 'linkage and exchange' models discussed in the KTE literature (Lomas 2007). In the survey, RDA evaluation officers put forward the potential for more informal (potentially voluntary) internal and self-evaluation processes for further 'constructive' feedback into delivery.

Project level evaluation in its existing form was acknowledged not only for evidencing directly attributable project level impacts, but also as a means of shaping delivery via the inclusion of formative (process) elements to address specific project needs and

⁹⁹ YF: 58%

leading to the generation of tacit knowledge. The roles of interim and ongoing evaluation were also found to facilitate remedial action to be taken if necessary. Formal mechanisms to embed this learning and assist organisational change were identified:

Interim project evaluations were very successful at informing the future delivery of projects. In some cases consultants went to the trouble of writing an implementation plan for the recommendations of the evaluation. (RDA evaluation officer [56])

More informal mechanisms were also revealed, whereby inclusion in the evaluation process influenced project managers' decision making behaviour:

I think the evaluation process, while by no means perfect, had a lot going for it. It made project managers think about what their projects were meant to achieve and whether they achieved them. (RDA project delivery/performance officer [39])

I'm never certain how much of the learning in the detailed final reports went on to impact future delivery at an operational level. But I'm certain it did in some way impact the way people approached their projects. (RDA project delivery/performance officer [12])

However, it was found that an evidence base built upon project level evaluation can be piecemeal and fail to draw out key organisational lessons. Drawing once again on Huber's terminology (2006), the RDA approach to evaluation did not result in a 'co-ordinated learning process', as knowledge was rarely diffused beyond the boundaries of individual project level evaluations:

I think the individual project interim evaluations were successful at what they did, but this shared learning should have been expanded out to a similar group of projects to share best practice. (RDA evaluation officer [56])

To summarise this section on the use of evaluation evidence within the RDAs, it was found that the uptake of evidence into policy processes was messy and complex, involving both formal and informal mechanisms. Evaluation was perceived to have a greater influence on demonstrating accountability than on enhancing the effectiveness of RDA policy making. It could be inferred that this predisposed evaluation evidence uptake in practice at all decision making levels. For RDA strategic decision making, it was found that evidence had less influence than other factors such as political

strategies, previous practice, ideas and values. This aligns with findings from Wells (2007) discussed in Chapter 2.

Instead, evaluation evidence was identified to have often been used as ‘window dressing’ (Huber 2006) to justify investment decisions that had already been taken. The analysis then suggests that the subordinate role of evaluation evidence within RDA strategic investment prioritisation was reflected within RDA project appraisal processes. It was found that incorporation of evaluation evidence into investment appraisal was limited to the minimum required for a purely ‘formal exercise’ (Huber 2006) and viewed as part of a ‘tick box exercise.’ The extent to which tacit knowledge gained from evaluation then influenced delivery was found to be directly linked to the level of engagement that delivery partners and RDA personnel had with the evaluation process. It was identified that learning was rarely diffused beyond the boundaries of individual project level evaluations, however, to enable a ‘co-ordinated learning process’ (Huber 2006).

8.7 The generation and communication of RDA evaluation evidence

The survey aimed to further explore both the generation of RDA evaluation evidence and the interactive process of communication between commissioners, producers, and users of evaluation evidence within the RDAs. This process of ‘Knowledge Transfer and Exchange’ (KTE) has been identified in the literature as not only disseminating timely, useful evidence-based research findings to decision makers (and others who use research), but also actively engaging such users in the research process to increase the relevance of studies (Mitton et al. 2007). Given that the policy making processes of the RDAs involved a range of discrete but interlinked activities, it was found that the KTE of RDA evaluation evidence was not a single process at the organisational level, but a multitude of parallel and successive processes (Rutter, Hawkins and Pankhurst 2013).

8.7.1 The generation of RDA evaluation and the role of ‘knowledge brokers’

In the survey, the highest rated capability of RDA evaluation reported by all respondents was that it ‘contributed to the evidence base’ (94%, 76)¹⁰⁰, and thus supported the development of regional intelligence. It was acknowledged that a key achievement of RDA evaluation processes was:

¹⁰⁰ YF: 92%

To get so many evaluations completed to reasonable standards in a short period of time. (External evaluator [67])

A general strengthening of RDA evaluation practice over time was found, which correlated with the RDAs maturing as delivery organisations:

There is little doubt in my mind that evaluation practice moved on significantly over this period. (RDA evaluation officer [53])

In general the [evaluation] processes worked well - they improved and became more consistent over time as RDAs became more experienced. (External evaluator [78])

A central theme of the survey was the 'knowledge brokering' role provided by evaluation personnel in terms of connecting and acting as an intermediary between external evaluators and RDA research and policy teams. Ward, House and Hammer (2009, p268) define knowledge brokers as "the interface between the worlds of researchers and decision makers, they are seen as the human force behind knowledge transfer, finding, assessing and interpreting evidence, facilitating interaction and identifying emerging research questions." This was reflected in the survey with respondents highlighting the need for evaluation personnel to possess not only specific technical and tacit knowledge, but also softer skills to facilitate interpersonal communication and collaboration:

The [evaluation] team were always approachable, knowledgeable, constructive, helpful and regarded as experts in their field across the organisation. (RDA economic appraisal officer [9])

The team had an excellent reputation both inside and outside the agency for skills, knowledge and effectiveness. (RDA strategy officer [59])

A trend of 'professionalisation' in the conduct of RDA evaluation in terms of planning for, commissioning and managing evaluations was identified. For instance, most respondents agreed that RDA evaluation processes: 'ensured evaluation met the guidelines of the RDA' (91%, 73)¹⁰¹; 'planned and prepared for evaluation' (86%, 68)¹⁰²; and 'ensured evaluation was carried out robustly' (79%, 63)¹⁰³.

¹⁰¹ YF:92%

¹⁰² YF:94%

¹⁰³ YF:79%

A great deal of time and energy was invested in the evaluation process, through procuring the most suitably qualified evaluation consultants for the project being evaluated, to setting up robust steering groups to direct the evaluation process, and constructive final meetings to present/clarify results/findings. (RDA Delivery/performance officer [12])

Respondents described a range of written and face-to face support that RDA evaluation officers used as a medium of exchange. In terms of planning for evaluation, respondents noted: evaluation officers writing and giving feedback on evaluation plans; providing guidance notes; delivering training on evaluation; and offering general support for project managers. In terms of commissioning and managing evaluations, respondents noted: the assignment of individuals dedicated to liaise between policy makers and external evaluators during the commissioning process; an improvement in the formulation of invitations to tender (ITTs); assembling and facilitating steering groups which involved representative stakeholders; the use of evaluation panels of consultants for commissioning; professionalisation of the working relationships with contractors including adopting fair practices for timescales and budgets; the participation of RDA evaluation officers in inception and steering group meetings; and the timely reviewing and feedback given to evaluation outputs. Overall, the data suggested that evaluation officers provided both a participatory and a quality assurance role for RDA evaluation. It has also been suggested in the literature that knowledge brokers may be one way of increasing successful dissemination practice (Armstrong et al. 2007).

8.7.2 The communication of RDA evaluation

Active dissemination and sharing of evaluation findings were identified as essential components for the evaluation strategies of the RDAs. This aligns with the finding of others, whereby passive dissemination has been acknowledged as ineffective (Kerner 2006; Grimshaw et al. 2006). It was raised by survey respondents that different types of information and communication styles are needed for different audiences. This supports Mitton et al. (2007) who highlight the need for innovative and targeted dissemination methods. It was identified in the survey that research in summary format, using simple language and highlighting targeted messages, is seen to be preferable and more likely to be taken up by policy makers:

Studies were often highly technical and the varied nature of RDA staff roles meant that key conclusions need to be presented simply so that lessons could be learnt without the technical detail. (RDA economic appraisal officer [10])

Respondents conveyed that robust, technical evaluation reports are necessary for detailing theoretical frameworks, methodological procedures and limitations. Yet, a balance needs to be struck against providing evidence that is distilled, easy to understand and that is relevant to decision making using policy appropriate terminology. In particular, the following methods were proposed to improve dissemination: distilling targeted lessons into summary documents; using face-to-face interactive meetings with key staff members and stakeholders; facilitating knowledge sharing seminar series events; and exploring online resources such as social media and data visualisation opportunities. Printed lengthy evaluation reports were reported to have less impact:

Final reports are rarely read in full by anyone other than the client project manager yet consultants spend days and weeks writing them. One page summaries, e-shots and newsletter summaries should be used much more to stimulate interest. Obviously a final report is obligatory, but it should be accompanied by more user-friendly outputs to prevent the evaluation dying a death in someone's inbox. (External evaluator [70])

Overall, the survey identified that RDA dissemination had not been effectively planned, resourced or evaluated to enable a coordinated learning process. 39% (31)¹⁰⁴ of all respondents agreed that RDA evaluation processes 'disseminated the results of evaluation within the RDA,' 35% (28)¹⁰⁵ agreed that they 'ensured learning and development from evaluation findings' and only 21% (17)¹⁰⁶ agreed that they 'disseminated the results of evaluation to external audiences.' This was seen as a missed opportunity, particularly by RDA evaluation officers themselves:

We should have recognised that dissemination of findings and the application of findings to action was about 50% of what we should have been doing. (RDA evaluation officer [56])

I would say that we had a lot of work to do on dissemination - we could have contributed to academic journals, put articles into trade journals... we could

¹⁰⁴ YF: 33%

¹⁰⁵ YF: 29%

¹⁰⁶ YF: 21%

have spoken at conferences, we could have led on strategy development - but we failed. (RDA evaluation officer [53])

The lack of analysing and synthesising evaluation evidence was identified as a key area of weakness:

There was a lack of analysis of evaluation and therefore dissemination and learning. (RDA project delivery/performance officer [19])

I'm sure there was lots of useful information contained in the various reports commissioned, but I'm not certain whether the overall learning from all this was combined in a meaningful and concise way... (RDA project delivery/performance officer [12])

Sharing lessons across the RDA network was also perceived to be a missed opportunity so that the RDAs could plan for, commission, and manage evaluation using shared best practice. It was suggested that more joint-working could have enabled: the influencing of evaluation guidance and practices from central government; working on jointly addressing problems with evaluation methodology/data; and cross-RDA commissioning of evaluations to consider wider evaluation research questions and to meet evaluation knowledge gaps.

Instead of BIS guidance advising the RDAs on planning and evaluating knowledge transfer functions, it was frequently cited by respondents that the demands of the Impact Evaluation Framework (IEF) and national RDA impact evaluation exercise displaced evaluation activities which could have impacted on organisational learning:

BIS expectations were a major challenge - resources had to be refocused on producing IEF compliant evaluations rather than evaluations that were useful for the organisation. (RDA evaluation officer [54])

The evaluation team had limited resources which were misdirected to achieving compliance with experimental frameworks from both BIS and the EU. These frameworks also demanded a minimum level of RDA spend to be covered by the evaluations. This led to little resources spent on analysing, disseminating and using the findings from the research. We were in effect reacting to the targets set by BIS rather than gaining information that was useful to the region. (RDA evaluation officer [56])

Van Der Knaap's work (2006) is relevant here as he argues that a focus on performance measures (applied here to consideration of the IEF) may inhibit the facilitation of dialogue, learning and responsive evaluation. Given that it was identified as a central

theme in the qualitative data, the influence of the IEF and the national RDA impact evaluation are explored further in the following section.

To summarise this section focused on the generation and communication of RDA evaluation evidence, a central theme of the survey was the ‘knowledge brokering’ role provided by RDA evaluation personnel in terms of connecting and acting as an intermediary between external evaluators and RDA research and policy teams. It was found that RDA evaluation officers provided both a collaborative and a quality assurance role for the generation of RDA evaluation evidence. A trend of ‘professionalisation’ in the conduct of RDA evaluation practice in terms of planning for, commissioning and managing evaluations was identified. However, it was found that there was a lack of a mandate for the structured dissemination and application of RDA evaluation findings.

8.8 Barriers and facilitators

By investigating the generation, communication and utilisation of RDA evaluation evidence, contextual factors (barriers and facilitators) for the uptake and use of evaluation evidence in practice were revealed.

8.8.1 A regulatory framework

The greatest perceived challenge faced by evaluation personnel in the survey was ‘changing guidelines and expectations from BIS’ (70%, 57).¹⁰⁷ A central theme in the qualitative data was the lack of a regulatory framework, and thus an organisational process, for the RDAs to supply and utilise evaluation evidence. In terms of the generation of evidence, a central theme was that evaluation had not been planned for and resourced from the start of the RDA’s existence:

The evaluation team... should have received more thought in the early stages of the RDA. (RDA project delivery/performance officer [34])

Linking researchers with users in the early stages has been identified in the KTE literature as a facilitator for the uptake of research into policy and practice (Ward, House and Hammer 2009; Greenhalgh, Howick and Maskrey 2014). Indeed, the methodological implications of the lack of an evaluation framework at the

¹⁰⁷ YF: 67%

establishment of the RDAs, and the challenges with then applying a common evaluation framework (the IEF) seven years into the RDA's operation were found to be significant (to be discussed in section 8.8.4). Trying to design an evaluation system once an organisation is up and running was also found to have implications on the demand side in terms of the perceived legitimacy of evaluation functions:

The IEF requirements provided the necessary focus internally to undertake evaluation... (RDA evaluation officer [49])

A higher profile and more 'authority' would have helped. (RDA Strategy officer [62])

As the design of RDA policy processes had not required evaluation evidence to be utilised systematically, this had implications for the 'pull' of evaluation evidence:

The work of the [evaluation] teams was largely prompted by the requirements of the BIS/PwC work and hence the evaluations were not as well embedded in the decision making procedures of the RDAs as they should have been. (External evaluator [67])

Respondents noted that the IEF and national RDA impact evaluation exercise initially gave 'weighting' and a sense of credibility to evaluation processes, raising the profile of evaluation and, in some RDAs, leading to the instigation of evaluation strategies, programmes of evaluation, evaluation officer roles being created and resources set aside for evaluation. Yet it was also noted that central government expectations of monitoring and evaluation also changed over time, which created uncertainty. Indeed, a complex array of performance management mechanisms was used to assess the successes of the RDAs (section 7.2). The manifestation of this central demand for evidence within the RDAs was articulated in the survey by one RDA evaluation officer:

A key operational challenge for the evaluation team was to disseminate valuable data that had real meaning. Frequently our evaluation team manager would be required to send ad hoc quantitative data to Government and rarely was [he/she] able/allowed to provide a brief description of what the data actually meant and its limitations, i.e. place it in a qualitative setting. (RDA evaluation officer [56])

This identified demand for "ad hoc quantitative data" in the above quotation aligns with the work of Sutcliffe and Court (2005, p3) who found that policy makers may demand a "limited range of 'top-end' evidence."

8.8.2 Political backing

When reviewing the utilisation and influence of evaluation evidence on strategic decision making, project appraisal and delivery it was revealed that other factors were of greater significance (see section 8.6). Organisational commitment to using evidence in decision making as set out in RDA strategic documents, signifying the formal acknowledgement of the role of evaluation, was reflected upon by respondents:

I think the lack of recognition of the importance of evaluation in both strategic and corporate documentation prevented the full benefits from being realised... (RDA strategy officer [66])

Less formal mechanisms for assisting the use of evaluation evidence were also identified. In particular, senior management support for RDA evaluation was identified to be important to 'cascade down' both to internal personnel and external partners:

Senior exec level commitment massively eased the challenges of persuading operational teams to engage positively with evaluation. (RDA evaluation officer [55])

The senior team played lip service to evaluation and whilst evaluation could have been viewed as essential to the evidence based policy making described in strategy documents the reality was at odds with this. (RDA strategy officer [63])

I was never really sure what senior management at [the RDA] thought about evaluation. Perhaps if they had been more vocal in their support both to [RDA] staff and our project delivery partners then it may have smoothed some problems over. (RDA project delivery/performance officer [12])

It was found that lack of encouragement by senior management to use evaluation evidence led to the perception by RDA project managers that evaluation was just an additional 'hoop to jump through' given high workloads and limited time:

Project managers sometimes seemed to view evaluation as being rather burdensome and of little relevance to their work. (RDA strategy officer [59])

Project Managers often saw evaluation as an add-on and came to the evaluation team late to plan or undertake an evaluation. Therefore it was very difficult for the team to effectively plan its own time accordingly. (RDA evaluation officer [56])

It was also observed that delivery partners and some project managers perceived evaluation primarily as a means of demonstrating accountability (i.e. as an audit function), rather than as a tool to shape delivery effectiveness:

No one (project managers or partners) like evaluation and as "a necessary evil" it does not get the priority it could. (RDA project delivery/performance officer [18])

Sometimes it was a nightmare. Contractors didn't really understand why external consultants were brought in to evaluate and they were really wary of them. (RDA project delivery/performance officer [12])

Respondents suggested that evaluations gained 'buy in' when they were undertaken for a purpose which aligned with the vested interests of delivery partners such as: to feed into decisions to secure additional investment; to showcase directly attributable project-level impacts; or when the evaluation encompassed their own research agenda:

When there was no obvious or direct link between the evaluation and a specific element of further investment, it was very hard to engage with partners. This meant that most evaluations lacked support and lacked interest, meaning that they were ultimately poor value for money. This is a pity, because the quality of the evaluations themselves was high. (RDA economic appraisal officer [3])

The term "culture" was frequently used (n=18) to describe the general receptivity to learning from evaluation evidence within the RDAs, and political backing from the 'top' was identified to be a key factor to encourage an evaluation culture:

The [evaluation] team's ability...was significantly hampered by the general ambivalence to evaluation shown by the Chief Executive and Directors and strategy team. (RDA strategy officer [63])

There was not that much interest in what our findings were! (RDA economic appraisal officer [3])

The culture within [the RDA] was not conducive to using evaluation evidence. (RDA strategy officer [63]).

In turn, respondents reflected upon strategies for increasing organisational buy-in for evaluation. The potential for individuals within senior management to champion evaluation and 'sell the benefits' was suggested by several respondents (Appendix 6).

Percy-Smith et al. (2002) also identified the need for 'championing from the top' in their work on the role of evidence based policy in Local Authorities. Similarly, Baumbusch et al. (2008) argue that if researchers become 'credible messengers' and decision makers become 'research champions', issues can be reframed and language be employed that is accessible to both researchers and practitioners.

Overall, it is interesting to note that the role of politics generated much (generally negative) discussion across the survey whereby politicisation of the process was observed to be problematic. This leads back to the theoretical discussion presented in Chapter 2 in that there was little reflection within the survey of the political and value laden nature of evidence. This seems to ignore the realities of decision making within the regional policy context.

8.8.3 Analytical skill

The ability of evaluation users to access, interpret and apply evaluation evidence to decision making was also identified to be important in the survey. Although the survey data suggested that over time awareness and understanding of evaluation increased at all levels in the RDAs, a deficit of analytical skill was still reported by respondents:

Evaluation... was not central to decision making largely because it was not understood. (RDA strategy officer [63])

The potential for capacity building was suggested, mainly in the form of educating and developing the skills of decision makers, via evaluation personnel delivering training and disseminating guidance materials. Indeed the need for decision makers, in particular senior management, to have the ability to assess both the potential and limitations of evaluation outputs was emphasised:

Support for, and a broad understanding of, evaluation is crucial at the top of the organisation to create a culture of evaluation. (RDA evaluation officer [49])

Interestingly, there was a lack of reflection in the survey data of the potential for professional development within evaluation teams to build knowledge transfer and communication skills. This aligns with the work of Ward, House and Hammer (2009) who point out that most papers on capacity building focus on developing the skills of decision makers, not researchers.

8.8.4 Data quality and availability

Issues with data quality and availability were identified to be key factors constraining the utilisation and impact of evaluation evidence within the RDAs. The lack of a common evaluation framework when the RDAs were established was once again reflected upon:

Evaluation should have been considered at the outset - establishing a framework to inform monitoring approaches thereby enabling a robust evaluation to be undertaken... (External evaluator [79])

Evaluations would have been more robust and informative had they been carried out systematically and comprehensively over a longer period of time and as an integral part of the investment cycle. (External evaluator [67])

In particular, the lack of a standardised approach to gathering evidence on key evaluation parameters leading to issues with data comparability and availability was a prevalent theme:

There could have been a more standardised approach to reporting and methods within RDAs to ensure comparability across findings. (External evaluator [75])

[Evaluation should have been undertaken] within a more universal framework and with a common series of outcomes/outputs in mind. The lack of the latter proved fatal. (Central Government officer [82])

I think working out a consistent model for what should be measured and how the information should be collected would have been helpful. (RDA evaluation officer [57])

This finding is supported by the paper ‘How to raise the bar on impact evaluation’ by Chadwick, Tyler and Warnock (2013). Respondents also perceived that it was a missed opportunity to better integrate data that monitored spend and outputs relating to beneficiaries of interventions, with emerging evaluation data:

[There] should be a dynamic process of collaboration [between monitoring and evaluation] to not only ensure delivery but to measure impact. (RDA project delivery/performance officer [38])

Formal evaluation and quantitative monitoring should have been more closely aligned. There was a complete disconnect between the very contractual/quantitative monitoring of the project... and the qualitative learning coming out of the evaluation process. (RDA project delivery/performance officer [12])

The need for agreement on the data to be collected at the beginning of projects and programmes to feed into decision making was emphasised. Respondents also underlined that there should be a contractual obligation for delivery partners to support and contribute to the evaluation processes, including agreement on the beneficiary data to be collected and reported upon:

For the most part, primary data was collected at the end of the project and often businesses did not recall in detail the assistance provided - it would have been better to carry out on-going monitoring for evaluation. (RDA evaluation officer [48])

Due to the lack of such measures it was found that there were resulting issues with both the quality and availability of monitoring and evaluation data:

If the quality and quantity of both evaluation AND monitoring data [had] been improved, it might have been possible to set up systems to make performance... more visible to decision makers. (RDA evaluation officer [58])

Many RDA evaluations, particularly quantitative assessment are of poor quality or of variable quality, which makes it difficult to synthesise. (RDA evaluation officer [46])

It was also stressed by RDA evaluation officers that it can be difficult, or inappropriate, to manipulate data to try and make it 'fit' into decision-relevant formats (such as benchmarks) retrospectively:

There were dangers... because the benchmark data were not derived from a process that was designed to provide benchmark data. (RDA evaluation officer [58])

Decision makers like evaluation to provide simple 'answers' - such as return on investment measures. In reality it is dangerous to base decisions on these - because of methodological limitations (e.g. not costing certain sorts of benefits and inconsistencies in treatment) and changing contexts. (RDA evaluation officer [54])

The survey data revealed that such methodological deliberations within evaluation were perceived by some decision makers as hindering the progression of developing robust, policy relevant evaluation outputs to meet policy makers' needs:

[Evaluation was] too academic focused and not enough focus on the real end point of evaluation which to me is its practical application to improve project delivery and investment prioritisation. (RDA strategy officer [63])

Indeed the RDA national impact evaluation exercise did ultimately rely on aggregated, quantitative evidence such as return on investment measures, despite issues with data quality and relevance. Ultimately, central government directly used evaluation evidence in the form of cost effectiveness ratios (i.e. £1 spend to £'s regional GVA created) for the RDA national impact evaluation exercise. Data were reported for each RDA and across the RDA network in the PWC report (PWC 2009a; PWC 2009b) and then disseminated by BIS.

Fundamentally, however, the IEF and national RDA impact evaluation were criticised by many respondents for being methodologically flawed. A major issue was that by the time a common evaluation framework (the IEF) was published, RDA structures, strategies and evaluation practice varied considerably across the RDA network. Respondents identified that there were therefore significant challenges with applying consistent methods and approaches to impact evaluation going forward and with retrospectively evaluating past activity using such an approach.

The main perceived limitation of applying the IEF was the weaknesses in the guidance surrounding the valuation of impacts (i.e. not costing certain sorts of benefits and inconsistencies in treatment). It was frequently highlighted that there is an ongoing challenge to define outcome indicators capable of practical measurement, and at relevant spatial levels, to assess the longer-term and wider impacts of policy interventions. In particular, the relative inability of current evaluation methodology to capture the impacts of tourism, public realm, skills and economic inclusion projects was noted. There was a sentiment that this undermined confidence in using evaluation evidence, and in particular GVA benchmarks, for future policy making as there was an inaccurate reflection of the potential impact of certain activities. Overall there was a general perception that the IEF methodology was thus incomplete and took a 'reductionist' approach focused on cost per job:

The IEF approach had limitations in assessing overall impact by drilling this down to cost per job. (RDA evaluation officer [49])

One of the weaknesses has been an inability to demonstrate the added value and spillover of regeneration and social investment. Anything that could be done in that field that HMT will accept would be invaluable. (Central Government officer [82])

The IEF and national RDA impact evaluation were also criticised for being prescriptive and mechanistic by several respondents, who particularly stressed the importance of addressing the context of individual interventions when undertaking regional policy evaluation:

IEF is grossly mechanistic in its focus upon all the theoretical steps in the gross-net adjustment, most of which are incapable of empirical research or involve application of standard assumptions. (External evaluator [73])

Context is the big issue - context, mechanism and outcomes are rarely differentiated. (RDA evaluation officer [54])

This aligns with findings of Polverari and Bachtler (2004) who note that inflexible evaluation frameworks can be counterproductive and yield evaluation techniques which can actually provide little insight into the real impact of regional policies. It was found that respondents also perceived that the IEF failed to embed an evaluation framework which collated evaluation evidence which was relevant to policy and practice:

The highly restrictive (and experimental) impact evaluation framework (IEF) set out by BIS meant a great deal of resources were spent collecting data that was not helpful at project level and was neither used at a strategic decision making level. (RDA evaluation officer [56])

[There was] more focus on trying to get evaluations and numbers that fitted the narrow definitions adopted by PWC and a complete disinterest in learning from evaluation. (RDA evaluation officer [51])

Overall, it was expressed by respondents that a technocratic understanding of evaluation was inbuilt through the audit approach methodology taken by PWC for the RDA national impact evaluation, with an excessive focus on monitoring and administrative practices rather than on outcomes and impacts. This finding is also supported by Chadwick, Tyler and Warnock (2013). Reflecting upon this, RDA evaluation officers noted that this placed focus on IEF compliant evaluation evidence and away from organisational learning and other types of evidence including tacit knowledge:

[There is a need to ensure the] clarity of evaluation objectives and purpose i.e. is the evaluation to produce a number for ranking or to evaluate more holistically our impact and draw out learning? (RDA evaluation officer [51])

The PWC work was an expensive numbers exercise. (RDA evaluation officer [51])

[The RDA's] focus was on learning and improving rather than counting. (RDA evaluation officer [51])

The survey data revealed the perception that the application of the IEF and the national RDA impact evaluation exercise ultimately produced work which had limited credibility with policy makers and which generated scepticism towards evaluation outputs. Huber (2006) argues that evaluation used as 'window dressing' gives a veneer of credibility to policy processes externally, whilst undermining evaluation practices internally. On one hand, there was a sentiment in the survey data that RDA evaluation findings and processes were undermined:

In retrospect, publication of IEF produced an industry and "experts" overnight... (External evaluator [73])

Evaluation will always be the poor relation to other more pressing needs, particularly given that... few believe the findings. (RDA evaluation officer [53])

On the other hand, there was a sentiment that that the RDAs had ultimately failed to utilise evaluation evidence to enhance the effectiveness of policy making:

[There is a] danger of evaluation being seen as something you do because you have to rather than want to. Consequences would be poor projects which don't deliver benefits, and not able to demonstrate adequately effectiveness - this has been crucial in the demise of RDAs! (RDA project delivery/performance officer [31])

An inability to make the case for the RDAs in terms of added value doomed them ultimately. (Central Government officer [82])

Nonetheless, it could be argued that RDA evaluation practice operated within 'bounded rationality.' There was no reference in the survey data of the desire to establish panels of beneficiaries (individuals or firms) for longitudinal analysis and considerations of quasi-experimental or RCT approaches were absent from the debate to start with, given the lack of an evaluation framework encouraging such study designs. A focus had remained on the use of certain evaluation building blocks including project beneficiary surveys, secondary data analysis and the occasional use of

models. There was also no mention in the survey of the need for a more robust peer-review process for evaluation outputs or the need for RDA evaluation practice to be 'evidence based', drawing from external (academic) research developments and programme evaluation outputs akin to NICE HTA appraisal processes in health policy (see Chapter 4). Overall, a strategic approach had not been taken to generate, communicate and utilise evaluation evidence within RDA policy processes.

To summarise this section focused on the contextual factors that influenced the generation, communication and use of evaluation evidence, a crucial finding is that a common evaluation framework had not been planned for and resourced from the start of the RDAs' existence. Accordingly, evidence was not gathered on key evaluation parameters to assess effectiveness at a national level (gross outputs and key elements of additionality, such as leakage, displacement and substitution) or to gather process evidence on the effectiveness and efficiency of a regional institutional framework compared to a more centralised/localised approach. It was found that the publishing of the IEF guidance and national RDA impact evaluation exercise instead embodied a mechanistic and managerialist understanding of evaluation that ultimately undermined the credibility of evaluation practice.

By investigating the generation, communication and utilisation of evidence, contextual factors (barriers and facilitators) were revealed for the uptake and use of evaluation evidence in practice. Barriers identified included: the lack of a regulatory framework and organisational processes leading to the 'pull' and structured use of evaluation evidence; political backing and support from the 'top' to encourage a culture of evaluation; guidance on planning and evaluating knowledge transfer functions; analytical skill and understanding of the potential and limitations of evaluation evidence; and significant issues with the quality, availability and relevance of monitoring and evaluation data. Facilitators identified included: involving users and evaluators early so that evaluation strategies can be established at the beginning of the lifetime of an organisation; using innovative and targeted dissemination methods; championing evaluation 'from the top'; providing opportunities for continuing professional development; and employing the use of knowledge management systems to support a structured dissemination strategy.

8.9 A potential role for knowledge translation tools

This section is set aside from the rest of the narrative above because it sets up the foundation for the following chapter. It was identified in the survey that targeted, decision-relevant evaluation outputs were more likely to be taken up by policy makers. Yet a number of issues relating to the access and relevance of RDA evaluation evidence were identified in the survey, potentially constraining dissemination. Firstly, the issue of the sheer volume of evidence requiring critical appraisal by decision makers was identified:

Information was lost in the waves of constant information fired at project managers. (RDA project delivery/performance officer [21])

The quantity of information will put off many people from even looking at the evaluation reports. (RDA evaluation officer [52])

Secondly, an issue with accessing and navigating RDA evaluation evidence was identified:

Evaluation reports should be more easily accessible internally and externally, e.g. reports posted on-line. (Delivery/Performance RDA officer [25])

I don't feel that findings from evaluations were freely available in order to help develop projects, especially cross-team and cross-directorate. (RDA project delivery/performance officer [11])

This highlights a potential opportunity for the use of knowledge management strategies and compiling evidence into a single authoritative source to package, translate and share RDA evaluation lessons (Ward, House and Hammer 2009). Finally, once again, the lack of analysis and evidence synthesis and the need to package quantitative evaluation outputs (what works) with qualitative outputs focused on learning (why does something work) were raised. Of particular interest is that both evaluation users and evaluation personnel emphasised the need for a more systematic approach and identified a potential role for the use of web-based repositories and databases:

I think having a more systematic approach to the exploitation of findings would have helped. (RDA evaluation officer [58])

[The RDA needed] the assembling of evaluation findings in a way which could be readily accessible i.e. some sort of interrogatable database. (RDA project delivery/performance officer [40])

[Need for] the ability to search and access data/information easily and remotely. (RDA project delivery/performance officer [40])

This insight provides a foundation for exploration of the potential use of a knowledge translation tool as investigated in the following chapter.

8.10 Conclusion

This chapter has contributed in a number of ways to address the second research question. It was found that in the RDAs other factors besides evidence, such as political strategies, previous practice, ideas and values, had a greater influence on policy decision making. This politicisation of the process generated much, generally negative, discussion across the survey responses. At first sight, and drawing upon the theoretical background presented in Chapter 2, this could be interpreted as respondents failing to acknowledge the realities of the complex decision making environment and the political nature of evidence. However, it is demonstrated in the chapter that evidence use was described as allowing for symbolic and ‘cherry-picked’ evidence to be used, rather than evidence being incorporated into policy making systematically. This leads back to the thread of discussion throughout the thesis about the appropriate use of evidence.

It was also found that, although an EBPM approach was embodied within regional policy making principles, evaluation was used as a tool for demonstrating accountability, rather than enhancing the effectiveness of RDA policy making through organisational learning. The publication of the IEF did provide a common approach to evaluation (albeit seven years into the operation of the RDAs); however, this chapter demonstrates that the demand for certain types of knowledge led to a central pull for evaluation that embodied a managerialist and mechanistic approach to EBPM. Overall, this empirical study of RDA policy processes affirms Sanderson’s (2006) theory of ‘instrumental bounded rationality’ discussed in Chapter 2.

By investigating the generation, communication and utilisation of evidence, contextual factors (barriers and facilitators) were revealed for the uptake and use of evaluation evidence in practice. It could be presumed that clarifying these factors is important for

planning well designed evaluation processes. In particular, a potential role for the use of knowledge translation tools was identified within the survey. This is investigated in further detail in the following chapter.

Using a Knowledge Tool to Extend an EBM Approach to Regional Policy Investment Prioritisation: A Critical Analysis

9.1 Introduction

This chapter addresses the final research question: what are the potential opportunities and barriers to using a knowledge tool to extend an EBM approach to regional policy investment prioritisation? This is explored through analysis of the perspectives of an expert stakeholder group and builds upon the groundwork provided in Chapters 4-8. The purpose is two-fold. Firstly, by extending an EBM approach for investment prioritisation to the regional policy context it is possible to, once again, draw upon the key differences across the sectors in terms of generating and using evaluation evidence. In this chapter the implications arising from differences in data quality and availability will be viewed through the lens of constructing a knowledge translation tool in practice. Secondly, it is possible to explore the use of a knowledge translation tool, in this case the use of a decision support tool, as a strategy to support an EBRPM approach and potentially increase the utilisation of evaluation evidence.

An online workshop was conducted with an expert stakeholder group to capture the perspectives of personnel engaged in RDA and local policy evaluation. Responses were elicited from senior commissioners, producers and users of evaluation evidence across the policy cycle. The theoretical background for the chapter is initially reflected upon, drawing upon the literature review. Then the characteristics of the workshop participants are discussed. The rest of the chapter is then structured by the themes that emerged from the literature review (Chapter 2), from constructing the decision tool (Chapter 3) and the workshop data including: the case of RDA investment prioritisation and the potential role for decision support; perspectives on using a knowledge tool to support an EBRPM approach; reflections on the construction of the knowledge tool; and perspectives on developing the knowledge tool.

9.2 Theoretical background

This part of the empirical research predominantly focuses on debates around the role of knowledge translation tools and decision support. Therefore the unabridged discussion of the theoretical background for this chapter is provided in Chapter 2 (predominantly section 2.6).

It was found in Chapter 4 that in the health sector, the establishment of NICE provided an institutional process, political backing and a legislative framework to incorporate evidence into strategic investment prioritisation decision making and, although open to strong criticism, this enabled a process of stakeholder engagement, peer-review and the setting of clear guidelines to evidence producers. In contrast, the RDAs were financed by national government public funds via the creation of a 'single pot' of RDA funding, where there was a high degree of flexibility, enabling funding to be targeted towards the investments the RDAs prioritised for each region (Chapter 5).

It was found in Chapter 8 that evaluation evidence had limited influence on formal mechanisms of investment prioritisation decision making at the strategic level within the RDAs and was sometimes cherry-picked or used symbolically to support decisions that had already been made. If evaluation evidence was systematically used, this was through the use of GVA benchmarks and it was contested whether or not this was an appropriate use of evidence given the issues with data quality, relevance and comparability. There was some evidence, however, of evaluation evidence being communicated and utilised through less formal mechanisms, potentially shaping strategic decision making behavior. Another relevant finding in Chapter 8 was that communication of RDA evaluation evidence was lacking and unsystematic. Knowledge management was identified as a potential strategy to support the structured dissemination of RDA evaluation evidence, providing access to research and other information in a single authoritative source.

Reflecting upon the above findings and drawing upon the groundwork so far (Chapters 4-8), it could be assumed that extending an EBM approach for investment prioritisation to the regional policy context would, once again, reveal key differences across the sectors in terms of generating and using evaluation evidence. In this chapter this analysis has been undertaken through the lens of constructing a knowledge

translation tool in practice. It has been hypothesised that implications arising from differences across the health and regional policy sectors, in terms of data quality and availability to underpin a decision tool, will be revealed. In addition, drawing upon the EBPM/KTE literature review (Chapter 2), it has been hypothesised that use of such a decision support tool has the potential to support an EBPM approach and to increase the utilisation of evidence.

In general, the methodology and subject matter for this chapter are relatively novel. Reflecting on the knowledge management literature, Driessen, Huijsen and Grootveld (2007) note that papers have been written on the use of knowledge mapping tools for instance, yet they note that not many papers have been written on the actual construction of such tools or on how to embed knowledge tools into organisational processes. The scope for this chapter, however, is firmly focused on the applicability issues with extending a NICE approach to RDA investment prioritisation and on the potential opportunities and barriers to the use of decision tools to support EBRPM, rather than on detailed consideration of the application of knowledge management theory and the programming details of the decision support tool.

9.3 Who were the workshop participants?

A total of 19¹⁰⁸ senior policy makers and analysts were recruited from three case study organisations including: former East Midlands Development Agency (EMDA) officers; former One North East (ONE) officers; and central government officers from the Department of Business, Innovation and Skills (BIS); and the Department for Communities and Local Government (CLG). A decision was made to not include Yorkshire Forward (YF) officers in the sample given that a prototype of the decision tool presented in the workshop was developed at YF which would have likely led to response bias.

As shown in Table 16, the expert stakeholder group that participated in the workshop included 3 EMDA officers (16%), 8 ONE officers (42%) and 8 central government officers (42%). The sample was roughly evenly split between RDA (11, 58%) and central

¹⁰⁸ One participant from ONE only contributed to open forums and did not respond to the scenarios. Due to a technical error, only 17 responses were recorded for scenario 1, with 18 responses recorded for scenario 2. See Chapter 3 for further details.

government (8, 42%) officers.¹⁰⁹ Given the key finding in Chapter 8 that senior management support was essential for the development of an evaluation culture, participants were recruited to the workshop from senior positions and included (former) Chief Economists, Assistant Directors, Programme Managers, Research and Evaluation Managers and Strategy and Policy Managers. Respondents were targeted if they were directly involved in the use of evaluation evidence for decision making within the RDAs or in current decision making processes for Local Growth initiatives within central government.

Table 16: Demographics of online workshop respondents

	Number	Percentage of sample
Respondent's organisation		
RDA:	11	58%
• EMDA	(3)	(16%)
• ONE	(8)	(42%)
Central government	8	42%

Self-completion surveys (questionnaires) were embedded into the online workshop and used closed questions to yield quantitative data and free text comments (open questions) to enable non-standardised responses. In addition, open forums with prompt questions enabled participants to post free text comments for open discussion. Detailed responses were given to the qualitative questions with a total word count across the surveys and open forum of over 9000 words. Therefore the workshop has provided a rich source of quantitative and qualitative data.

9.4 Presentation of the findings

For the quantitative data, the responses presented are for the total sample, and a number and a percentage are reported. For the qualitative data, quotations are presented denoting the departmental sub-group and identification number of the respondent.

For further information: survey items and respondents are presented in Appendix 7; the quantitative data are presented in Appendix 14; the qualitative data are presented in Appendix 15; and further consideration of the online workshop method and limitations are presented in section 3.8.

¹⁰⁹ Appendix 13 presents the full list of participants.

9.5 Investment prioritisation without the use of a decision support tool (scenario 1)

9.5.1 Strategic investment prioritisation in the RDAs

It was found in Chapter 8 that evaluation evidence had limited influence on formal mechanisms of investment prioritisation decision making at the strategic level within the RDAs. On one hand, it was found that communication of RDA evaluation evidence was lacking and unsystematic. On the other hand, it was found that in the RDAs other factors besides evidence, such as political strategies, previous practice, ideas and values, had a greater influence on policy decision making. Politicisation of the process was perceived to be problematic because it led to symbolic and cherry-picked evidence use. The inverse of this is that there was a concern over the importance of systematic and unbiased evidence use within the policy process. Overall, a need was identified for targeted, systematic dissemination and the generation of decision-relevant evaluation outputs.

To answer the question of ‘what does systematic, unbiased, decision-relevant evaluation evidence look like?’ the online workshop focused in on the (hypothetical) decision point of strategic investment prioritisation between RDA projects and programmes.¹¹⁰ When considering the workshop data, an underlying assumption was immediately apparent. There was a perception that directly comparing the potential costs and benefits of RDA investments was only appropriate within programmes where projects were of a similar nature. It was not acknowledged that such decision making occurs, indirectly, when allocating budgets strategically between programmes:

Ranking projects is always tricky, comparing apples and pears. Hence allocating budgets to programme areas, and then sub-programmes, and ranking within those sub programmes is easier - comparing apples with apples. However, the choice then has to be made in allocating levels of budget to different programmes and sub-programmes. (ONE officer [11])

¹¹⁰ In reality, as demonstrated in Chapter 8, RDA economic appraisal work was usually focused towards major projects undergoing government appraisal, rather than for all RDA investments. Given that most RDAs took a project (rather than a programme) approach to delivery, and given that the RDAs often prioritised investment on an ad hoc basis rather than through formal funding rounds, projects and programmes weren’t often directly compared against each other.

Individual projects cannot be judged against each other in my view unless they are similar in nature - this needs to be part of a strategic decision that says we need to spend XX on infrastructure, YY on skills and ZZ on business support to achieve a certain amount of GVA. (ONE officer [16])

[I] would look at projects within individual programme/sub themes, to compare similar projects to each other. (ONE officer [11])

9.5.2 A hypothetical investment decision making scenario

The workshop aimed to investigate the subjective manner in which participants individually processed decision-relevant information. The unabridged discussion of the methodology for this chapter is provided in Chapter 3 (predominantly sections 3.7 and 3.8). To give an overview, in the workshop for scenario 1 participants were asked to rank 10 economic development interventions across a range of programmes (i.e. policy areas) against each other within a revised budget (reflecting a budget cut of total investment equating to 50%).

Participants were presented with limited headline data on the 10 projects to guide and inform their decision making and to be utilised as they saw fit. Quantitative data were presented on expected costs and benefits and qualitative data were presented in the form of an overview of the project and an overview of pragmatic delivery issues. This dummy data was based upon a review of RDA data collated to underpin the decision support tool (section 3.7). An example of the information provided to participants for scenario 1 is presented in Table 17.¹¹¹

¹¹¹ See Appendix 9 for the full data presented to participants.

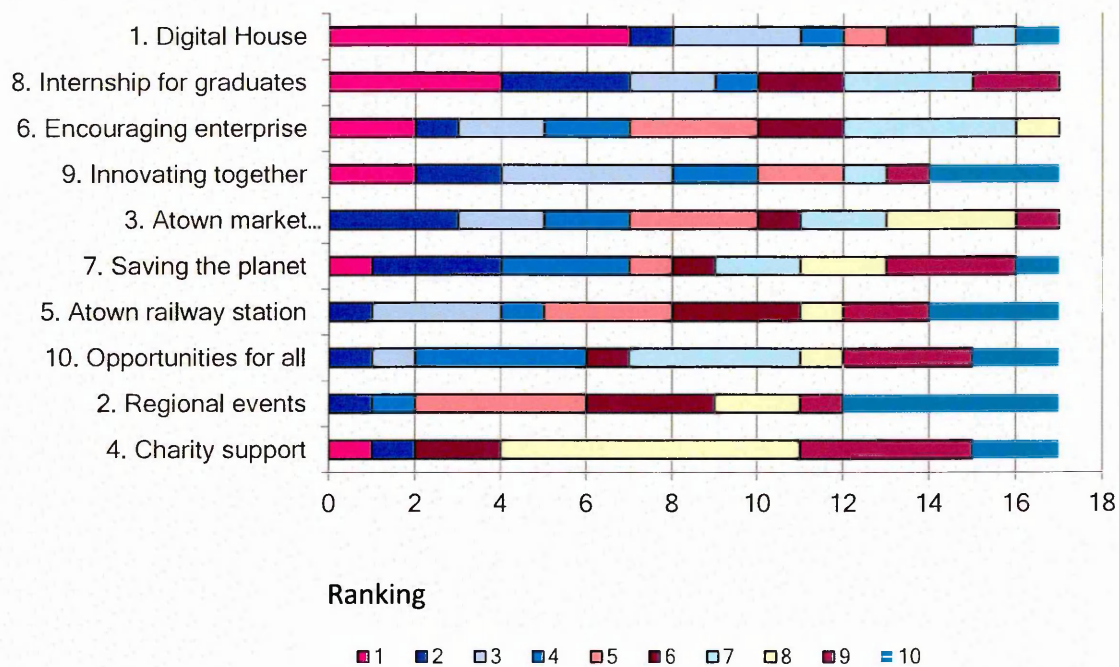
Table 17: Example information presented to participants for scenario 1

Example projects		
	Digital House	Atown market regeneration
Project description	Building purchased to develop into high quality office accommodation for digital companies	Redevelopment of town market complex into a theatre, café and town council offices
Local Authority	Barnsley District	North Lincolnshire
Project investment in £s (2011-12)	£150,000	£300,000
Total GVA in £s (for investment 2011-12)	£3,075,000	£4,230,000
Jobs created	100	5
Businesses created	0	5
Businesses supported	0	20
People assisted to gain employment	0	0
People assisted in skills	0	0
CO2 reduced (tonnes)	0	0
Project manager's notes	Took 5 years to acquire working with the local authority. Currently an eye-sore in the centre of the town. A lot of negative publicity in the local press about how long it has taken to begin development.	Most of this project has been carried out. This remaining budget is for completion and doing an evaluation to look at the impact of the investment.

Detailed discussion could follow consideration of Table 17, reflecting upon the relevance and credibility of each of the categories of information presented. However, the main point to be made is that when participants' ranked the projects against each other for scenario 1 and submitted their survey responses, analysis of the quantitative data revealed that the vote distribution (i.e. demonstrating how each participant ranked the interventions) was highly complex, even for a limited number of interventions. As shown in Figure 19, there was no clear consensus in the data about the order participants thought the projects should be ranked.¹¹²

¹¹² See Appendix 14 for the calculated overall rank and score.

Figure 19: Distribution of votes



Several workshop participants reported that they aimed to use a cost benefit type approach to rank the interventions for scenario 1:

I prioritised jobs created and GVA and investment costs to generate a benefit cost ratio which I feel is key to decision making. (ONE officer [7])

I used an approach that looked at the ratio of costs to deliverables. I took account of evaluation evidence for similar kinds of projects with which I am familiar. And in the context of the current economic context, prioritised job creation and business starts. (EMDA officer [1])

My reaction to scenario 1 was to build into the spreadsheets a cost benefit model with weights for the benefits given. (Central Government officer [8])

So, if a purely analytical approach had been taken to objectively weigh up the merits of the data categories to come to a decision, then Figure 19 suggests that participants gave different weightings to the categorises of data presented and/or used different methods when ranking the interventions. This would perhaps reflect that the analytical ability of each of the participants varied.

However, the manner in which participants reported they individually processed the information, discloses highly complex decision making behaviours:

I gave each one a score for 'priority' based on my gut feel of the fit with my personal RDA priority list, another score for outputs, and a third for GVA: £1 ratio (noting...skepticism about the figures). I then ranked each of these scores and calculated an average weighting. Finally, I then fiddled the weightings until I got a result I was happy with! (ONE officer [11]. Forum comment)

I ruled out some (ranked 5-10) due to lack of rationale... or expensive and no exit strategy... In real world projects would need a business case, with economists looking at economic case (I'm not an economist - as you can probably tell!). Subject to that I ranked the others on - policy imperative..., evidence of (likely) success...; vfm and in an area of known need for the economy (skills, enterprise). I looked at cost per job as a v. rough indicator of vfm. (Central Government officer [18]. Forum comment)

The quotations above demonstrate that, not only are the objective merits of the decision-relevant categories of information considered, but individual agency factors (such as instinct and judgement), experiential knowledge of previous practice (such as deliverability and organisational objectives) and other factors (such as political strategies) are taken into account. Reflecting back on Sanderson's work (2003, p339-340) in Chapter 2, the differing roles of 'episteme' (theoretical academic and research knowledge/evidence) and 'techne' (instrumental professional and institutional experience), as well as 'phronesis' (intrinsic virtues embodied in human practices during decision making) can be identified.

The balance between these factors for decision making was considered by some participants:

It needs to be a careful balance between evidence and common sense. A steady middle ground would improve replicability and consistency. (Central Government officer [17])

It could be inferred then, that the lack of consensus in ranking the data in Figure 19 could also suggest that each of the decision makers brought varied, wide-ranging background experience and knowledge to the process alongside individual values and ideas.

Drawing all of this together, it is interesting to note that participants frequently requested a broader range of information (beyond the distilled information presented in the scenario) to robustly inform decision making processes:

I would have needed more information - set out in the business case and based on evaluation of similar projects. (Central Government officer [15])

Therefore, a key finding of scenario 1 is that decision makers expected to be able to undertake a complex synthesis of information cognitively; critically appraising evidence of varying relevance, quality and comparability, combining this with tacit knowledge and employing the use of subjective judgement.

9.5.3 An identified role for knowledge translation tools

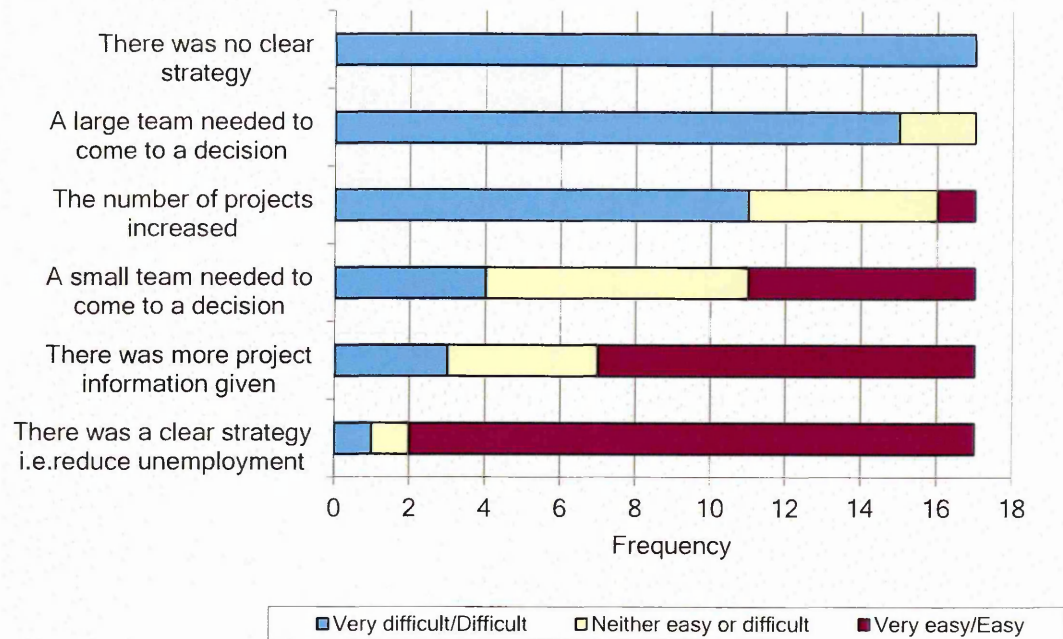
Potential constraints to individuals' analytically processing information were explored through the workshop. The reason for this line of investigation was two-fold. Firstly, the application of tacit knowledge and subjective judgement were identified to be important in decision making (section 9.5.2). However, McCaughey and Bruning (2010) argue that 'heuristic' bias errors, defined as errors during intuitive cognitive processing (i.e. when using mental 'rules of thumb' and 'gut feeling responses') may have implications for the rationality assumptions of evidence based decision making (Chapter 2). This may suggest a rationale for the use of decision support to enable systematic and unbiased evidence use.

Secondly, Borenstein et al. (2009) argue that the subjective manner in which individuals' process decision-relevant information may become compromised as the decision making context becomes more complex. They note (Borenstein et al. 2009, pxxii):

While a reviewer may be able to synthesise data from a few studies in their head, the process becomes difficult and eventually untenable as the number of studies increases.

In order to explore the factors which add to the complexity of decision making processes, participants were asked to consider how easy or difficult it would have been to undertake the investment prioritisation exercise in scenario 1 if the decision making context were changed (Figure 20).

Figure 20: The context of the decision making process



In Figure 20, participants reported that having an overall strategy agreed for the decision making process was the most important factor impacting upon decision makers individually processing decision-relevant information. It was found that this cognitive process was perceived to become more difficult if there was an increase in the number of interventions to compare (where evidence provided may be disparate in relevance, quality and comparability) and as the group decision making environment becomes more complex, involving multiple stakeholder preferences. This finding was also reflected in the qualitative data:

In my prioritisation I used an underlying clear strategy based on economics. Without this it would have been very difficult. With a large team it is more difficult to agree this underlying strategy and ensure everybody is pulling in the same direction. (Central Government officer [2])

There was a tendency with RDA investment decisions - that I saw - for the volume of evidence to stand in the way of clear decision making. (ONE officer [12])

The complexity and volume of information could itself become a problem for decision makers. (EMDA officer [9])

However, once again Figure 20 demonstrates that participants perceived a broader range of information was needed, beyond the distilled information presented in the scenario, to inform their decision making. This supports the earlier finding that that

there is a balance to be made between presenting enough detail to add context to distilled information, against the need to surface key messages (Chapter 8). It could be argued that, past a certain level of information being provided, diminishing returns set in hampering decision makers to surface the key messages:

The process must be sensitive to the limited capacity of decision makers to process information. (EMDA officer [1])

Reflecting upon the above discussion, it could be argued that these factors identify an analytical rationale for the use of decision support tool.

In addition, and concurring with the findings of the online survey in Chapter 8, participants in the workshop also discussed their experience of policy processes allowing for strategic and cherry-picked evidence use. This was, once again, identified to be problematic:

Personally I would have chosen those projects which deliver the highest forecast GVA and which deliver GVA over a long period of time (i.e. physical redevelopment). In reality, political considerations - both national and local - would inevitably mean that 'pet' projects would be championed and commissioned. (EMDA officer [5])

Political realities can mean that the most justifiable decision may be different from the best decision. For this reason it is often tricky to accurately record decision making processes. (ONE officer [16])

Therefore it could be argued that the use of a decision support tool may enable policy decisions to be informed drawing on rigorous, systematic, and un-biased evidence, rather than allowing for strategic, cherry-picked and biased evidence to be used.

9.6 Using a decision tool to support an EBRPM approach (scenario 2)

A simple, prototype decision support tool was constructed. It was used as a means to elicit research participants' perspectives on the use of the tool for decision making in the workshop. How the decision support tool was actually constructed is described in the methodology chapter (see section 3.7). To give an overview, the decision support tool was designed to support RDA strategic investment prioritisation decision making and to allocate budgets across programmes and projects. It was programmed to enable a change in the total budget available (i.e. a change of the resource allocation threshold). The tool drew upon a knowledge base of evidence drawn from across the

RDA directorates and employed a decision logic so that users would be able to simulate aspects of a decision scenario alongside summarising the evidence base (Chapter 2). An EBM approach was programmed into the underlying model of the tool to reveal practical analytical and data access implications of extending an EBM approach to the regional policy context. The decision tool was then presented to workshop participants in an introduction video and participants were asked to consider prioritising investments (i.e. ranking projects) with the use of a decision tool in scenario 2 (see section 3.8).

9.6.1 Receptivity to the decision support tool for investment prioritisation

Investigating the potential opportunities for the use of the decision support tool required exploration of its purpose within the policy process. This leads back to the discussion in Chapter 2 about the aims of the communication and utilisation of evaluation evidence. For instance, Armstrong and Wells (2006b, p265) note that there is sharp disagreement over how evaluation findings should be applied, with those who feel that evaluators should impose evidence informed change if needed at one end of the spectrum and those who feel that evaluators do not have this mandate and instead “should facilitate actors to reach a deeper understanding of what they are doing” at the other.

Participants were asked to consider if they would use such a support tool for decision making, reflecting on their previous or current roles. Although no participants stated they definitely would not use a similar decision tool, 22% (4) stated they were ‘unlikely to use’ such a tool. The majority of participants answered positively with 67% (12) and 11% (2) stating they would be ‘likely to use’ and ‘definitely would use’ such a tool respectively.

It was identified in the workshop that participants particularly responded positively to the potential use of decision tools as a strategy to facilitate evidence use.

Quantitatively, 100% (18) of the participants agreed that the decision tool facilitates the use of available data. It was also explored within the workshop whether the use of a decision tool could support EBRPM by facilitating decision making to be both accountable and effective (linking back to the principles for regional policy discussed in section 7.2). Although 78% (14) of participants agreed that the decision tool would

facilitate decisions to be replicable and 72% (13) agreed that it would facilitate decisions to be transparent, only 33% (6) agreed that the use of the decision tool would enable better outcomes, such as an improvement in GVA or outputs.

These findings were reflected in the qualitative data. Receptivity to the use of the decision tool in scenario 2 compared to scenario 1 (where no decision tool was provided) was frequently positive:

Much better - more evidence-based, transparent and data-driven. (EMDA officer [5])

We should do much more of this in decision making. (EMDA officer [9])

In particular, it was once again emphasised that using a decision tool was an effective strategy for increasing the systematic utilisation of evidence:

I... did work up a similar model at ONE. It's an excellent way of actually using real evaluation data to help inform future investment decisions. (ONE officer [7])

This sort of tool is a great way of using evaluation evidence intensively and effectively. (ONE officer [7])

The tool provides much more information in a consistent format... I think the systematic use of evaluation evidence/benchmarks is a strength of the model. (EMDA officer [1])

The potential opportunity for the decision tool to present data in a form which enables data cleaning, synthesis and analysis was identified:

It would ensure consistency in figure work and allow various financial options to be considered with relative ease, providing there is accuracy in input data. (Central Government officer [3])

Presentation of the data is much more transparent and easier to slice and dice. (ONE officer [7])

A very useful tool to aid decision making, by holding and ordering a lot of information in a single place as well as ensuring the process is transparent so all involved in the process are clear on the criteria. (ONE officer [6])

In particular, as somewhat expected given the discussion in section 9.5.3, the potential opportunity for the decision tool to address constraints to individuals' analytical processing of information was identified:

The tool helps inform the process, particularly with large numbers of projects. (ONE officer [10])

And to counter cherry picked and unsystematic evidence use:

While no tool will capture all qualitative data, this nonetheless takes us forward. And it is an improvement on many of the finger in the air/ personality based investment decisions made in many organisations. (ONE officer [7])

Many of the quotations above suggest that use of a decision tool was perceived to provide opportunities to increase the transparency, and therefore accountability, of decision making:

[It is] very useful to be able to show a robust method for prioritising spend. (ONE officer [6])

However, when considering the role of a decision tool to shape policy decisions, the responses were markedly less positive in tone.

9.6.2 The role of a decision support tool to shape policy decisions

When considering a deterministic role for the decision support tool to 'enable better outcomes,' by directly providing evidence based decisions or dictating policy, it was found that participants countered that this approach would ignore the complexity of decision making in reality:

There is a risk that this process can become over scientific. (EMDA officer [9])

Plugging figures into a sausage machine such as this discourages true, deeper analysis. (ONE officer [11])

Danger of oversimplifying the process. (Central Government officer [4])

It was also countered that such an approach would ignore the political nature of evidence and decision making:

You can't get away from the political angle, and maybe you shouldn't try to. Robust quantitative models are great, but you have to allow people to have

their say otherwise the process looks remote and done by eggheads in ivory towers. (ONE officer [12])

There is a balance to be struck between what is pragmatic (i.e. public and political considerations) and the outcome of quantitative or other analysis... in the real world, you have to be aware of the views of politicians and the electorate, and that could change the decisions made. (ONE officer [12])

However, a potential role was identified for the decision tool to *inform* decision making and potentially shape policy within the wider process rather than for dictating policy. This aligns with Ottoson's (2009) conceptualisation of evidence being used for 'bottom up change' rather than 'top down change'.

It was underlined that, in reality, decisions are rarely made by individuals and are often made collaboratively by a team of individuals across an organisation, or even across several. The importance of negotiation and peer challenge was discussed by participants as a way to ensure that interventions were realistic and met strategic priorities. Some participants perceived that the tool could be a useful starting point to enable an initial analysis of the data before negotiation processes. However, others perceived that the tool could be used as an integral component of the negotiation process, to structure debate:

It seems to me to form the basis of a dialogue - you can [show] stakeholders what the model says, and which projects it prioritises, and then have a discussion about the pros and cons. Much of this discussion will be of a political nature. (ONE officer [12])

Overall, the above discussion on receptivity to the tool has surfaced the potential opportunities for the use of a decision tool to support an EBRPM approach. The next section now goes on to consider the barriers.

9.7 The construction of a decision support tool

A simple, prototype decision support tool was constructed. It was used as a means to elicit research participants' perspectives on the use of the tool for decision making in the workshop. How the decision support tool was actually constructed is described in the methodology chapter (section 3.7). It was found that the construction of a decision tool to support investment prioritisation within the regional policy context, underpinned by RDA evaluation and monitoring data, was highly problematic. This section draws upon the experience of constructing the decision tool (Chapter 3). It also

draws upon quotations from the online survey (Chapter 8) and the online workshop, focused on technical data quality problems.

9.7.1 Drawing upon a knowledge base

Knowledge mapping

When constructing the decision tool, it was found that the RDAs had not taken a holistic approach to integrate the collection of monitoring, evaluation and socioeconomic data into one authoritative source (also found in Chapter 8). Therefore, as discussed in Chapter 3, when constructing the decision tool a process of mapping internal and external data/knowledge sources within YF was undertaken, going beyond monitoring and evaluation data to link to other intelligence sources. This included brainstorming with key personnel in other teams supplying data, and undertaking unstructured interviews with decision makers about relevant decision scenarios and model outputs. This initial tool was underpinned by data collated from 'Artemis' (the RDA's Management Information System(MIS)), from a survey with project managers, the Chief Economist's Unit, the contracting team, the legal team and from published benchmark evaluation data. Given issues of data access and confidentiality, for the purposes of the research dummy data were developed to feed into a simplified version of the tool.

Data quality

When reviewing monitoring data from 'Artemis' (YF's Management Information System (MIS)) and from published benchmark evaluation data, significant issues with the quality of the evidence base were revealed, constraining the construction of a decision support tool. Likewise, workshop participants identified that the quality of RDA evaluation and monitoring data was a significant issue across the RDA network:

Any tool which is data driven inevitably lives or dies by the data which drives it. (EMDA officer [5])

From experience, a lot of the core data is a bit iffy. So the tool is only as good as the data that's put in. Screening the poor data would need to be a key task upfront. (ONE officer [7])

The model is clearly a useful tool, but is clearly highly dependent on the quality of data involved. Having worked on the evaluation framework both in the RDA's

and Government, that would be a key challenge. (Central Government officer [17])

This aligns with findings from the survey:

If the quality and quantity of both evaluation AND monitoring data been improved, it might have been possible to set up systems to make performance (including return on investment and social and environmental performance) within and across projects more visible to the strategic decision makers. (RDA evaluation officer [59])

These issues are investigated in further depth below.

The use of GVA

A significant issue when constructing the decision tool was the perceived quality of GVA data. Concerns surrounding the composition of the GVA measure and the robustness of the GVA benchmarks have been acknowledged (Chapter 5). Accordingly, when sourcing data for the decision tool, it was found that decision makers within YF placed emphasis on output data over GVA benchmark data. This experience of constructing the decision tool was supported by responses from EMDA, ONE and central government officers in the workshop. In the quantitative data it was found that GVA was assigned a much lower level of importance (47%, 8) than the core outputs of jobs created (82%, 14) and businesses created (82%, 14). A number of respondents questioned the credibility of GVA indicators:

I tended not to believe the GVA figures. Outputs, evaluation findings and knowledge about the delivery partner are stronger factors, but harder to do objectively. (EMDA officer [9]. Forum comment)

It was identified that not all RDA intervention types were well suited to analysis of their expected impact by GVA forecasts:

Idiosyncratic activities may have little evaluation evidence on which to base any benchmarks. (EMDA officer [5])

It was also identified that GVA benchmarks were not designed to be specific to each RDA:

Some of the assumptions are open to criticism. In particular, the multipliers from the PWC evaluation may be the best available, but they may vary from region to region or for other reasons. (ONE officer [12])

The relevance of cost per output data for decision making was also identified in the online survey findings (Chapter 8):

The flaw at [the RDA] was following through with hard metrics on cost per output. (RDA project delivery/performance officer [29]. Online survey)

As well as in the workshop:

[The decision tool] could give information in a much more tangible form of the various output/ outcome options which can be of more importance than the actual financial reduction to GVA. (Central Government officer [3])

Therefore, the decision tool was constructed so that users had the option to rank interventions against each other using the common currency 'cost per output' (instead of £ to GVA). However, regional policy interventions often have multiple strands of activity, multiple outputs and outcomes, quantitative and qualitative outcome measures and long-term outcomes (Chapter 6, Table 13). Thus there was a need to capture the complexity of projects within RDA MIS systems. In terms of constructing a decision tool it was therefore found that the YF monitoring system was not designed to produce cost per output data and that projects were assigned multiple outputs. To produce the prototype decision tool, a crude method employed was to interview senior managers to elicit the primary output (for instance, jobs created) for each programme area and to then survey project managers to elicit the proportion of investment that was to be spent by each project on activities to generate this primary output. This gave crude cost per output data.

Evaluation data problems

Other significant issues were found with the RDA evaluation data when constructing the decision tool. It was found that data on key elements of additionality (i.e. leakage, displacement and substitution) were not extracted from evaluation studies and stored within a single authoritative source for the RDA network.¹¹³ It was also found that

¹¹³ This was also a key finding of Chapter 8 where a need for knowledge management was identified.

there were issues with missing and inconsistent data within and across RDA evaluation studies.

The RDA national impact evaluation exercise had demonstrated that the IEF had been applied inconsistently across the RDA network and within RDAs (Chapter 5).

Accordingly, it was found that additionality data were missing and inconsistent within individual evaluations. For instance, some evaluations were not able to be IEF compliant and provide estimates of GVA impact (ONS 2011) and not only were methods for calculating GVA found to be different, reporting of GVA varied. For instance, it was ambiguous whether GVA to date, cumulative GVA and/or forecast GVA were reported, meaning that data were overlapping and inconsistent. In addition, a review of the 400 (plus) published evaluation reports from the RDA network on the Office of Project and Programme Advice and Training's (OffPAT) e-library repository (OffPAT 2012) also revealed that not all evaluations were published on OffPAT. For the construction of a decision tool these data quality problems led to issues with inconsistent aggregating and inconsistent timing of the data.

The synthesis of evaluation data was further hampered by evaluation data problems. Regional policy evaluation methodological guidelines did not go into prescriptive detail about how to apply sensitivity analysis methods to report the confidence level and confidence interval/margin of error achieved for survey results (if using a probability sample) (Chapter 5). Likewise, it was found that within individual evaluations, reporting of uncertainty was variable. For the construction of a decision tool this meant that approaches could not be applied such as the meta-analysis of evaluation studies that rely on the precision of the data and size of the study to be reported (Borenstein et al. 2009). The lack of the use of sensitivity analysis was also identified in the online survey (Chapter 8):

[Decision makers should be made] very aware of the limitations of the data used including the levels of confidence they could be expected to have in them. Then they might have felt more comfortable using evaluation to its potential in strategic decision making. (RDA evaluation officer [58]. Online survey)

And in the online workshop:

Need to take into account risk. Sensitivity analysis should be more common. (Central Government officer [4])

Overall, these data quality issues meant that there were single source and multi-source problems with aggregating evaluation data (Rahm and Do 2000). This is not surprising as it was found in Chapter 5 that the RDA national impact evaluation highlighted that, fundamentally, the quality and study designs for the evaluation varied greatly and it was not possible to aggregate the data to generate a robust analysis of the overall economic impact of the RDAs on the national economy as a whole (PWC 2009a; PWC 2009b).

The difficulty with aggregating evaluation data when the common evaluation framework (IEF) had only been published 7 years into the operation of the RDAs was discussed in Chapter 8. Issues with data aggregation were also reflected upon in the workshop:

[I] know from experience that evaluation data from different sources may be 'calibrated' differently. (ONE officer [10])

Monitoring data problems

The construction of a decision tool also required the utilisation of monitoring data. During the review of the YF monitoring system it was found that there were issues of redundant and duplicate data as well as issues of contradiction and optimism bias (i.e. missing, inconsistent and overlapping data). Reasons for this stated in the online survey included data entry errors:

To clarify – [regarding] the lack of monitoring data and its robustness. This was due to staff not inputting outputs onto the computer system in a timely fashion and by staff not being professional when they inputted it and essentially making it up at times. (RDA strategy officer [63]. Online survey)

The approach to monitoring was not consistent across all teams. (RDA delivery/performance officer [34]. Online survey)

As well as potential conflict over the categorisations in the technical guidance and monitoring system (i.e. leading to poor schema design):

The whole system of monitoring outputs was very complicated... and people's interpretation of the guidance on what constituted a verifiable output made monitoring really difficult. (RDA delivery/performance officer [12]. Online survey)

Within the workshop, responses from EMDA, ONE officers and central government officers also emphasised issues with the quality of RDA monitoring data. Thus the importance of cleansed data was identified:

Using cleansed data will provide a useful tool and could avoid costly mistakes going forward. (Central Government officer [3])

Underlying logic model

Overall, the review of data to construct the decision support tool demonstrated that the lack of an integrated monitoring and evaluation framework at the beginning of the lifetime of the RDAs (Chapter 8) led to the collection of data without an understanding of how such data were to be aggregated:

As for monitoring and evaluation data, the trick is to establish what you need to collect BEFORE you start commissioning projects. (EMDA officer [5]. Forum comment)

Linked to this, it was found that it was not clear how such data were connected to an underlying programme theory and that there was confusion over the identification of outputs, outcomes and impacts in the monitoring data. This is a theme that has run throughout this thesis. It was discussed in Chapter 6 that theory-based approaches aim to build understanding of the reasons for effectiveness and the circumstances under which results are likely to be replicated, thus going some way to account for the complexity of the context and contingency between variables. In the online survey a role for 'Realistic Evaluation' programme theory (Pawson and Tilley 1997) was identified:

Context is the big issue - context, mechanism and outcomes are rarely differentiated. (RDA evaluation officer [55]. Online survey)

Similarly in the workshop the use of logic models to map out the causal chain from inputs to outcomes and impact, and testing the underlying assumptions (Tavistock Institute 2003) was identified:

[There should be] more use of recognised logic chains. (Central Government officer [3])

Leading back to discussion in Chapter 6, of pertinence is that Sefton (2000) argues 'Theory of Change' models and standard economic approaches to evaluation both follow an input-outcome framework. Therefore he argues it is theoretically possible to examine both variations in costs and outcomes as well as taking into account context variables. The evaluator starts by defining long-term objectives and works backwards from the endpoint through the steps required to get there. Early stage or intermediate objectives are then established for each step, so that the programme can be evaluated, and if necessary modified, at any stage (Connell and Kubisch 1998). Long-term objectives are more likely to focus on outcome measures and shorter term objectives are more likely to focus on process measures. It could be envisaged this would require economic models of both the relationship between policy intervention and the outcome metrics and the relationship between socio-economic metrics and the outcome metrics.

Overall, it could be argued that an evaluation and monitoring framework established at the start of an organisation's existence may facilitate the collection of more appropriate data linked to an underlying programme theory. The development of a knowledge translation tool may then enable tangible understanding of how such data are to be aggregated.

9.7.2 Employing a decision logic

The decision tool was constructed so that it employed a decision logic, enabling users to be able to simulate aspects of a decision scenario alongside summarising the evidence base. The decision tool was constructed so that interventions could be sorted by a common currency (cost per output) with a threshold applied at the point when the RDA budget is exhausted (Chapter 3). This was based upon a simplified version of NICE decision making processes (Chapter 4) whereby QALYs are combined with the relative cost of treatment to form an 'Incremental Cost-Effectiveness Ratio' (ICER) (i.e. the cost per additional unit of health gain). This is then compared to an allocation threshold and the maximum acceptable ICERs are £20,000-£30,000 per QALY gained (NICE 2013b).¹¹⁴

¹¹⁴ However, it was demonstrated in Chapter 4 that the resource allocation threshold NICE employs is not directly related to the NHS budget.

The decision tool was constructed so that it ranked projects in order of those that had higher benefits compared to costs, and only took forward projects which fell under the (revised) resource allocation threshold. Cost per output was used to rank projects against each other within the model. Then the decision tool would equally split the total budget across programme types (i.e. skills, business assists, CO2 reduction, etc.). If the budget were reduced, then it would be the best performing projects within each programme that the decision tool would suggest should go forward and the total costs and benefits of these projects would be presented. This was a crude method and future programming would have needed to strengthen the allocation of the budget between intervention 'types,' mapping 'primary' outputs and outcomes more clearly onto potential future programme types.

The workshop highlighted the problems with the simple decision logic chosen for constructing the tool. A key issue was the valuation and direct comparison of outputs. Responses from scenario 1 revealed that workshop participants assigned differing importance to the various output categories:

My main concern was to focus on the core business of an RDA, and away from areas that are other organisations' responsibilities, such as CO2 reduction or skills. (ONE officer [11])

I would prioritise the regeneration and business creation schemes, and leave the less tangible programmes. Experience and hindsight suggests they are the scheme that got the greatest bang for your money. (Central Government officer [17])

Co2 measures are often underrated but shouldn't be. (ONE officer [7])

These quotations reflect differing preferences towards certain types of outputs and thus, policy areas. As described in section 9.5, such preferences may have been shaped by individual agency factors (such as instinct and judgement), experiential knowledge of previous practice (such as deliverability and organisational objectives) and other factors (such as political strategies).

Within scenario 2, when participants were presented with a decision support tool, it was emphasised that only limited user input had been programmed into the prototype decision tool. This effectively enabled a simple binary response (yes/no) to taking forward certain intervention types within the analysis (where primary output type was

used as a proxy to define the intervention type). Feedback from workshop participants strongly emphasised the need for a more nuanced approach to understand decision-maker's preferences, and convert this into an algebraic form for the programming of the model:

[There needs to be] stakeholder agreed criteria, scorings and weightings. (Central Government officer [8])

I think such a model is very useful but it would need modification and the decision maker should be able to add some additional knowledge into the decision making process. (Central Government officer [2])

Getting the decision makers to articulate their views would be the key to this exercise. Without this process it will be difficult to get a basis for decision that will bear scrutiny. (Central Government officer [8])

Overall, future programming may have enabled decision makers to place weightings (i.e. preferences) on the different programme areas (i.e. 20% of the budget to be allocated to CO2 reduction). How such weightings could be obtained to be fed into the decision tool is discussed in section 9.8.3 below.

9.8 Developing the decision support tool

A number of suggested improvements to the decision tool were raised through the workshop.

9.8.1 Developing the knowledge base: tacit knowledge

A central theme of the feedback on the prototype decision tool focused on the importance of considering non-technocratic and qualitative information:

I don't think that the model provides adequate coverage of more qualitative sources of information that are often very important in the context of investment decisions. (EMDA officer [1])

Although the need for robust, high quality quantitative data was once again emphasised:

Numbers are not the full story, and neither should they be. But as a first step they provide an excellent means of sorting the wheat from the chaff. (EMDA officer [5]. Forum comment)

Nonetheless, the need to understand the context of distilled quantitative data was once again underlined:

The reality is usually that you need to delve beneath the figures to understand the true impact of the project and the need for the funding. (ONE officer [16])

Participants fed back that the knowledge base, which the decision tool draws upon, would need to be expanded. Additional quantitative factors were suggested such as the persistence and timing of outputs and the future funding profile. The need to understand the socio-economic rationale for an intervention and the potential for greater data-linking to spatial data was identified. Additional qualitative factors put forward to be taken into consideration included: 'lessons learned' and best practice knowledge from evaluations; deliverability; risks associated with the project; Strategic Added Value (SAV) criteria; social, environmental and wider impacts; potential leverage; fit with government policy; and the 'political fallout' of terminating contracts:

By the end of [the RDA], we would have used an appraisal process that included use of evaluation evidence to inform decisions. However, more qualitative inputs such as views on deliverability, risks associated with the project and 'political' sensitivities were also taken into account. (EMDA officer [1])

[The tool] doesn't, as far as I can tell, allow for flexibility to deal with political or pragmatic issues. It also doesn't seem to factor in the power of negotiation with project partners - a project might not be worth going ahead with at the current cost, but at a reduced cost may well be worth it. (Central Government officer [15])

I would use [the tool] as part of the process, understanding there are quite often other factors which also need to be taken into account in decision making, i.e. practicalities, politics, partners, short term priorities, etc. (ONE officer [6])

The importance of experiential knowledge was once again emphasised:

The danger is that decision making is over-dependent on just cost-benefit analysis and does not provide enough possibilities for well-informed experts with experience of appraising and making investment decisions. (Central Government officer [4])

It must also provide for a balance of quantitative and qualitative inputs. Also it is important to acknowledge that no tool can make an investment decision - (subjective) judgment will still be required on the part of the decision maker - and this is as it should be! (EMDA officer [1])

And the importance of value judgements was once again raised:

Outputs [are] important, but gut instinct about the types of projects that are most successful is important but hard to measure. (EMDA officer [9])

The tool would have been helpful to inform the decision making process, but would have presumably only been part of the process. I don't think it can replicate the personal judgment of the project officer e.g. whether a project could be scaled back and still achieve sufficient outcomes. (Central Government officer [15])

While it is important to capture the information that scenario 2 provides, and important to use this sort of model to defend difficult decisions, the risk is that it produces outcomes that "feel" wrong. (EMDA officer [9])

It was strongly felt by some participants, however, that such non-technocratic, qualitative evidence could not be incorporated appropriately into decision tools:

[It is] impossible to create a model that incorporates qualitative data. (ONE officer [16])

Based upon such feedback, a middle ground could be proposed to develop a knowledge management repository, rather than a decision tool. As a counterpoint, however, Multi-Criteria Analysis (MCA) was suggested by some participants as a means of strengthening a decision support approach.

9.8.2 Multi Criteria Analysis¹¹⁵

A potential role was identified for the decision tool to be used as a basis to structure negotiation and consensus-building within collaborative decision making environments (section 9.6.2). MCA has been identified as providing a means of identifying, analysing and weighting stakeholder preferences alongside other forms of evidence (Dodgson et al. 2009). It could be conjectured that an extension to the model could therefore be to build in a MCA 'front end' to link stakeholder feedback on agreed assessment criteria through into the model weightings. Theoretically the importance of such

¹¹⁵ Dodgson et al. note (2009, p20) that "MCA establishes preferences between options by reference to an explicit set of objectives that the decision making body has identified, and for which it has established measurable criteria to assess the extent to which the objectives have been achieved." They add, "a key feature of MCA is its emphasis on the judgement of the decision making team... MCA can bring a degree of structure, analysis and openness to classes of decisions that lie beyond the practical reach of Cost Benefit Analysis" (Dodgson et al. 2009, p20).

preferences/weightings could be tested through ‘one-way sensitivity analysis,’ observing changes in the decision tool’s outputs as the inputs are changed. A decision support tool, with a MCA dimension, may have potentially been able to facilitate consideration of political and pragmatic factors and non-directly comparable information, such as the level of investment to be allocated between the disparate RDA programmes.

The potential use of MCA was identified by 3 workshop participants. One participant explained how the method would be applied for scenario 2:

Decide on the criteria for assessing bids, such as: rationale for government intervention, economic impact, jobs created (should be part of economic impact...), political priorities. Use simple multi-criteria analysis to rank bids and adjust using experience in assessing bids to come up with a list of suggested funding options. Give senior decision makers decision on final projects. (Central Government officer [4])

Another participant explained the technique in this extended quotation:

In real life and, if asked for advice, I would have recommended some form of MCD (multi-criteria decision making) technique. Such techniques, which are usually applied in a workshop setting and require good quality facilitation, aim to (i) explicitly tease out the criteria against which decisions are being made (ii) attempt to build some consensus and understanding about the trade-offs between different decision options in terms of these criteria (iii) place the type of information that former colleagues have identified (the details that might emerge in negotiation and project shaping) in a coherent framework. These techniques are ideal for choosing between a range of options particularly where resource constraints exist. They help identify the strengths and weaknesses of particular options and, through this, can lead to option improvement as a side-product. (ONE officer [19]. Forum quotation)

These extended quotations have been included to highlight that, if an MCA dimension was to be incorporated into the decision tool, this would require expert advice for the initial programming as well as ongoing technical support. Expert advice would also be needed to structure MCA workshops and facilitate the relationship between analysts and policy makers. A transdisciplinary approach to this would be needed given that MCA techniques have not been widely applied within the regional policy sector.

9.8.3 Updating data over time

Another key point identified in the workshop for the development of the decision tool was that it would need to be flexible and updated to account for changes in context and the emergence of new evidence:

The underlying assumptions would need to be regularly updated or re-considered. (EMDA officer [5])

To be reasonably accurate the data needs to be captured early in the projects life cycle and continually updated as retrospection can have its own flaws. (Central Government officer [3])

It was identified in the workshop that one of the key issues undermining the influence of evaluation evidence is the mismatch in the policy-research cycle. Therefore it was suggested that there is perceived value in generally ‘keeping on top of the evidence’, and taking a more regular, systematic approach to evaluation evidence synthesis. This aligns with findings from Rutter (2012) who explored the role of evidence and evaluation in policy making more widely across government.

Of interest, is that it was noted by participants that a systematic process for this could be to update the benchmark data in the decision support tool as and when new evaluation evidence was produced. This is demonstrated by this extended quote:

This requirement for evaluation should be across all projects, regardless of size and complexity, and should be based on a singular set of guidelines (so the same data is being collected for all projects). This means that at the end of a project it is possible to say what the actual benefits of the work have been (and using the concept of your model, these findings would then be incorporated into the benchmarks figures used to assess future projects). (EMDA officer [5])

It could be conjectured that an extension to the model could be to build in a Bayesian meta-analysis approach (Higgins, Thompson and Spiegelhalter 2009) to enable a formal mechanism for sequentially combining new evidence with the existing evidence base. Theoretically, an element of Bayesian Meta-Analysis could be programmed into the static look-up tables of such a decision support tool by the inclusion of ‘informed priors’ (potentially using ‘WinBUGS’ software, see Spiegelhalter et al. 2003). The aim would be to statistically combine Return on Investment (ROI) and ‘additionality’ data with new evidence as it is produced, with the posterior distribution calculated from one stage becoming the prior distribution for the next (potentially using cumulative

meta-analysis, see Lau, Schmid and Chalmers 1995). Such techniques could combine the precision but relatively small weighting of an individual evaluation's results on 'additionality' and expected GVA benchmarks, with existing benchmarks in the decision tool. The weighting given to an individual evaluation when combining it with the existing benchmark would depend on the confidence in the data.

Theoretically over time, the decision tool could become more powerful, and the predictions improve in certainty, as more data are obtained. Techniques such as probabilistic sensitivity analysis could then be conducted to calculate confidence intervals for key decision criteria such as total GVA and expected outputs using software such as 'Crystal Ball' (www.crystalball.com). Once again, a transdisciplinary approach to this would be needed given that Bayesian meta-analysis techniques have not been widely applied within the regional policy sector. However, this development to the decision tool is unlikely given that it would be dependent upon work on a 'common currency' (benchmarks) to have been undertaken/updated and for precision (uncertainty) to be reported in studies to be included in the analysis.

9.8.4 Programmer skill and the need for capacity building

The greatest concern raised in the workshop about the use of the decision tool was the role of hidden assumptions within the programming of the tool. Some participants dismissed the decision tool outright due to such inbuilt assumptions in the tool and others stressed that agreement on the inputs and programming of the model was vital:

Decision making is easier [with the tool] but may be over-reliant on the underlying data and assumptions. (Central Government officer [4])

Too many assumptions made, giving an illusion of objectivity whereas really the subjective decisions have been hidden in the assumptions built into the model. (ONE officer [11])

This finding highlights the role of communication between analysts and policy makers when developing such tools. The critical need to pilot-test the decision tool, and to establish if the options work as expected, was highlighted as a means of validating the tool and building confidence and trust with decision makers:

I would want to see it applied to some real life examples to judge its effectiveness. (Central Government officer [15])

Overall, it was stressed that decision makers would need to be fully aware of the limitations of the data and analysis:

While it gives an apparent objective comparison between projects, it is hard for non-experts to see how the numbers have been arrived at. (ONE officer [11])

[Needs to be] clear indication of the assumptions and evidence used to calculate the cost and benefits. (Central Government officer [4])

This once again highlights that the programming of such decision support tools would need to be based upon best-practice and to be undertaken by analysts with sufficient analytical skill. In addition, it was once again highlighted that capacity building would be needed:

The biggest problem which such a tool is that it would need to be both endorsed and ratified at the highest political level (whether it be at an internal, local, regional or national level). (EMDA officer [5]. Forum comment)

[The] tool isn't too difficult to understand but there would have to be a very good understanding, and buy-in, across the organisation so it had credibility. (ONE officer [6])

The public sector, alas, has many barriers which inhibit [decision support tools] use including significant cultural factors and a lack of analytical understanding. (ONE officer [19]. Forum quotation)

These quotations highlight that implementation of a decision tool would need to be facilitated by specialists with both sufficient analytical and communication skills to gain 'buy in.'

Overall, it is theoretically interesting to consider issues surrounding the construction and development of a decision support tool based upon an EBM approach and to consider how such a tool may be embedded within a regional policy organisation. However, the above discussion suggests that the use of a decision tool employing a knowledge base and decision logic is beyond reach within regional policy, for at least the short term, given problems with the quality and availability of the current evidence base. This insight leads on to the final point.

9.8.5 The role of a decision support tool to shape the generation of evidence

It is therefore interesting to consider what evidence based regional policy making may have 'looked like' if a national evaluation framework and decision support tool had

been formulated when the RDAs were established, ready to be populated with emerging evidence.

It could be theorised that the development of a knowledge translation tool, which is operationalised through an evaluation and monitoring framework from the start of an organisation's existence, may facilitate the collection of more appropriate, decision-relevant data linked to an underlying programme theory. It could be argued that this would enable tangible understanding of how data is to be aggregated, highlight any knowledge gaps and facilitate data-linking to other intelligence sources. It may also generate problem-driven evidence where it is already understood how such evidence is to be used for decision making. Indeed, McCaughey and Bruning (2010) note that the process of gathering information and clarifying the policy objectives of various stakeholders may in turn help to shape policy formation, delivery, monitoring and evaluation.

This long-term approach would permit concern for the rigour of evidence (according to the methodologies applied) and the importance of systematic and unbiased evidence to be communicated and integrated within the policy process. However, it also recognises that policy decisions ultimately remain based upon other factors besides evidence such as political strategies, previous practice, ideas and values.

Indeed, to return to the comparison with the health sector, the prototype decision tool was based upon a simplified version of NICE decision making processes so that interventions could be sorted by a 'common currency' and compared to an allocation threshold. In reality, NICE has identified that research evidence is not enough and that policy should be informed by evidence rather than based upon evidence. Although open to strong criticism, NICE uses bodies like Citizen Councils to undertake stakeholder consultation and help glean the views of the wider world (Chapter 4).

9.9 Conclusion

This chapter has contributed in a number of ways to address the final research question. It is revealed that decision makers expect to be able to undertake a complex synthesis of information cognitively; combining quantitative data alongside experiential knowledge of deliverability, organisational objectives and political ideas and employing the use of subjective judgement. When it came to ranking data within

an investment prioritisation decision making scenario, each of the participants took a different approach and assigned different weightings to the discrete categories of information. It could be inferred that this reflects not only that the analytical ability of each of the decision makers varied, but also that they brought wide-ranging background experience and knowledge to the process.

However, it was also found that there are constraints to the cognitive, analytical processing of information, which may provide a rationale for the use of decision support tools. Heuristic bias errors were discussed and it was found that the cognitive synthesis of information becomes compromised as there are increases in the amount of information provided (which may be disparate in relevance, quality and comparability) and as the group decision making environment becomes more complex, involving multiple stakeholder preferences.

A decision support tool was constructed and the process reflected upon within this chapter. It is shown that extending an EBM approach for investment prioritisation to the regional policy context has, once again, revealed key differences across the sectors in terms of generating and using evaluation evidence. However, undertaking this analysis by programming an EBM approach into the underlying model for a decision tool to support regional policy investment prioritisation, revealed practical, analytical, and data access implications.

It is therefore theorised that the development of a knowledge translation tool, which is operationalised through an evaluation and monitoring framework from the start of an organisation's existence, may facilitate the collection of more appropriate, decision-relevant data linked to an underlying programme theory. It was found that this may enable tangible understanding of how data is to be aggregated, highlight any knowledge gaps and facilitate data-linking to other intelligence sources. It was found that a knowledge translation tool would need to be flexible and updated to account for changes in context and the emergence of new evidence. It would also need to be programmed, supported and the relationship between researchers and policy makers facilitated by specialists with both sufficient analytical and communication skills.

This chapter demonstrates that the use of a knowledge translation tool dictating policy or providing 'decisions' ignores the complexity and political nature of decision making.

However, it could be argued that their construction and use may enable the more appropriate and systematic utilisation of evaluation evidence to shape policy decision making behaviour within the wider process.

Chapter 10

Conclusion

10.1 Introduction

This chapter outlines the overall significance, key findings and limitations of the research conducted. The research problem and gaps in the current knowledge base are considered and then the overarching research findings are described by synthesising and drawing out the key ideas from the preceding chapters to demonstrate the study's original contributions to knowledge. The pedagogical applications of the research, in terms of the transferability of the findings to other contexts and the practical and theoretical implications of the work for policy and practice, are then reflected upon alongside recommending further potential avenues of future research. To finish, the concluding remarks of the thesis are presented.

10.2 The research problem

This thesis tests the applicability of, and the factors which determine the success of, Evidence Based *Regional* Policy Making in England. The regional policy context provides a fascinating, and previously under-researched case study to explore the wider EBPM debates given the complexities associated with its multifaceted policy agenda, structures and mechanisms, alongside its inherently political character.

A key issue that defines the research problem is that although the academic literature is well established and becoming increasingly sophisticated, the practice of regional policy evaluation in England by the institutions charged with its implementation has not kept pace with this development. Less attention has been paid to the evaluation of UK regional policy instruments (i.e. non EU policy instruments) and to the processes of undertaking regional policy evaluation in practice. It is mainly departmental and professional reports that have reviewed the processes of regional policy evaluation in England (e.g. Cook et al. 2008; NAO 2010; ONS 2011; NAO 2013a) and there has been a call for lessons to be learnt from the RDA evaluation experience (Great Britain, Parliament, House of Commons 2010; Chadwick, Tyler and Warnock 2013; Garretsen et al. 2013). However, to the researcher's knowledge, one recent paper by Chadwick,

Tyler and Warnock (2013) stands alone in the academic literature examining the processes of impact evaluation within the Regional Development Agencies (RDAs). Thus, given the recent shift of focus for regional policy evaluation within the wider 'what works' agenda (Haynes et al. 2012; Cabinet Office 2013; BIS 2014b), the implications of extending an approach more closely aligned with Evidence Based Medicine (EBM) are relatively unknown. There is no single unifying account of EBPM; it is used in different ways across the policy and academic worlds (Wells 2007).

Evidence based *regional* policy making is the focus of this PhD. An interdisciplinary approach across the health and regional policy sectors has been taken. There were two major research aims:

Research aim 1: A dominant perspective within the EBPM literature has been to use EBM approaches as a yardstick against which wider social policy is assessed and to draw parallels between the practises of EBM and EBPM (Sefton 2000; Sefton et al. 2002; Sefton 2003; Dobrow, Goel and Upshur 2004; Cookson 2005; Somekh et al. 2005).

This research investigated how various types of evidence/knowledge are used across contexts and with different actors to understand what can be deduced about the generation, communication and use of regional policy evaluation evidence. Although there have been attempts to understand the utilisation of evaluation more widely within central (NAO 2010; Rutter 2012; BIS 2014b) and local government (Percy-Smith et al. 2002; Allen, Grace and Martin 2014), there has not been a study exploring the utilisation of evaluation or the extrapolation of an EBM approach to the regional policy context.

Research aim 2: The utilisation of knowledge translation tools and decision support tools as a mechanism to support the uptake of evidence is an emerging topic in the EBPM literature. Yet the decision support literature identifies a potential opportunity for simulating aspects of a decision scenario alongside summarising the evidence base (Gorry and Scott Morton 1971; Pervan and Arnott 2005; Eom and Kim 2006; and French, Maule, and Papamichail 2009), as well as the potential to address cognitive capabilities and biases (March 1978; March 1994; McCaughey and Bruning 2010).

This research critically analysed the role of a decision support tool to extend an EBM approach to regional policy investment prioritisation decision making.

10.3 The research questions and key findings

This section draws upon the key findings of all phases of the research to address each of the research questions in turn. Three key themes emerged when reflecting upon the significance and findings of the research: Evidence Based Regional Policy Making and the role of evaluation; the role of knowledge translation tools and decision support; and the research management insights gained from conducting the online workshop. Each of these themes is considered in turn.

10.3.1 EBRPM and the role of evaluation

This theme relates to the first research aim and to the first and second research questions.

Research question 1: *What are the epistemological and applicability challenges of extending an Evidence Based Medicine approach to regional policy evaluation?*

The method employed to investigate this research question was a comparative analysis of the methodological guidelines and central ‘pull’ for evaluation evidence across the health and regional policy sectors, drawing upon the academic and policy literature (Chapter 3).

Initially, a review of the Evidence Based Policy Making (EBPM), Knowledge Transfer and Exchange (KTE) and political science literature (Chapter 2) enabled the identification of three key cross-cutting debates: what kinds of evidence are used and the role of research credibility (‘what’); the issues surrounding the way in which evidence is incorporated into the policy making process (‘how’); and what are the other factors besides evidence which affect the way policy is made (‘other factors’). These overarching EBPM debates provided a conceptual framework to draw parallels across the policy sectors of health policy and Evidence Based Medicine (EBM) in Chapter 4 and Regional Policy and impact evaluation in Chapter 5. Mirroring this analysis highlighted key differences across the sectors in terms of generating evaluation evidence and using it for investment prioritisation. Then in Chapter 6 a direct comparison was made across the sectors and the challenges of extending an

EBM approach to regional policy evaluation were surfaced. Chapter 6 includes original analysis providing: a comparison of EBPM debates across EBM and regional policy evaluation (Table 11); an epistemological comparison (Table 12); and a methodological comparison (Table 13). The findings of such investigations revealed both potential opportunities and challenges to extending an EBM approach to regional policy in reality.

In terms of *potential opportunities*, comparison of the sectors highlighted that there has not been a formal 'evidence based movement' within regional policy equivalent to EBM within health policy. The analysis in Chapter 5 identified that evaluation evidence has not played a major role in regional policy development and showed that there is a gap between practice and academic thinking. In contrast, analysis in Chapter 4 revealed that within health policy, EBM has grown to become a large and powerful movement and its sphere of influence has expanded far beyond its origins in internal medicine. Therefore, there has been critical scrutiny and debate in the literature over the different types and credibility of evidence and the balance between individual policy maker expertise (organisational/practitioner knowledge), patient values and expectations (user knowledge), and external evidence (research/ policy community knowledge). In contrast within regional policy, it was found in Chapter 5 that less attention has been paid in the literature to evidence hierarchies and the role of research credibility. An expectation for how evidence was to be incorporated into policy making processes, and the division between central and regional agencies, has been somewhat unclear. Conversely, in health policy, Chapter 4 demonstrated that overall the establishment of NICE has provided an institutional process, political backing and a legislative framework to incorporate evidence into policy making. Although open to strong criticism, this has enabled a process of stakeholder engagement, peer-review and the setting of clear guidelines to evidence producers. The Institute has also created demand for economic evaluation evidence and the use of 'modelling', research synthesis and sensitivity analysis.

The analysis in Chapter 4 identified that the EBM rational choice model fails to engage with the political nature of decision making, however, meaning that the actual delivery of health care is likely to be somewhat different in practice. This was found to be particularly the case when moving beyond the micro to the macro decision making

level, where context becomes increasingly important. It can be deduced that the application of a “one size fits all” (Goodman 1999, p250) NICE (HTA type) evaluation model, which fails to take into account wider social values and interests, is therefore less reconcilable within the complex decision making environments for wider health and social care policy. Accordingly, consumers of evidence need to be cautious of the false sense of certainty which could be created by the NICE process and understand the limitations of the evidence base to prevent misled policy formulation (Bovaird 2014). On a similar vein, Chapter 5 demonstrated that the context of regional policy is very different, and in some ways is more complex than health. The complexities associated with regional policy’s multifaceted agenda, structures and mechanisms alongside its inherently political character transform the nature of the evaluation process. Regional policy evaluation has very different aims from evaluations conducted within the EBM paradigm and needs to address political issues and choices alongside intervention efficacy.

Drawing upon this foundation, it was found in Chapter 6 that a significant underlying debate for regional policy evaluation, which has certainly come to the fore recently, has focused on trust in the reliability of research findings. This has led to a call for the extension of approaches more closely aligned with EBM to wider social policy such as the use of RCTs, the establishment of a ‘NICE for social policy’ and the use of quasi-experimental approaches and economic evaluation (Haynes et al. 2012; Cabinet Office 2013; BIS 2014b).

By mirroring analysis of the generation and use of evaluation evidence across the health and regional policy contexts, it has been possible to highlight their differences to reveal how parallel EBM processes for evaluation and investment prioritisation are played out in different ways within each context. It was found that there are a number of implicit assumptions embedded within EBM and within EBRPM which have direct implications for an extension of an EBM approach to regional policy (and to wider health and social policy) and significant epistemological, methodological and practical implications were identified. Reflecting upon this, it is perhaps unsurprising that it was also found that claims that RCTs and quasi-experimental approaches can and should be applied within spatial policy (BIS 2014b) are already beginning to unravel in reality as the evaluations of new local growth initiatives are being commissioned. Likewise, the

WWG has faced significant challenges with applying systematic review and meta-analysis to the current spatial policy evidence base.

It could be concluded that, in some ways, there has been an opportunity missed in extending the EBM approach. While NICE, and by implication EBM, has a well-defined but relatively limited role (albeit contested at the margins) with respect to the whole of the NHS, in regional policy the potential scope for EBRPM is much wider (i.e. to include all project/programme activity). Even allowing for the difference in scale between the two domains, it is rather ironic that the study findings indicate that the potentially pervasive role in regional policy for EBPM following the NICE model is most unlikely to be implemented. Still, a more nuanced review of the vast literature on EBM, the methodologies employed and a greater understanding of the NICE process, alongside a greater appreciation of the political nature of decision making, could have provided a richer insight into the appropriate use of evidence within regional policy making. This stands in contrast to simply appealing to the increased uptake of particular forms of evidence.

Research question 2: *What factors influenced the generation, communication and use of evaluation evidence within the English RDAs?*

The method employed to investigate this research question was the development of an online survey, conducted with an expert stakeholder group involved in RDA regional policy evaluation (Chapter 3 and 8). Responses were elicited from eighty-one policy makers and analysts, including commissioners, producers, and users of evaluation evidence. The survey was designed to collect standardised quantitative and free text qualitative data. Detailed responses were given to the questions, with a total word count across the survey of over forty thousand words. This qualitative data provided rich detail and the quantitative data indicated the underlying trends.

The survey enabled exploration of the application of the regional policy evaluation guidelines and the central pull for evaluation evidence within the RDAs, which revealed debates around how evidence was incorporated into policy making processes. It was found that in the RDAs other factors besides evidence, such as political strategies, previous practice, ideas and values, had a greater influence on policy decision making. This politicisation of the process generated much, generally negative, discussion across

the survey responses. At first sight, and drawing upon the theoretical background presented in Chapter 2, this could be interpreted as respondents failing to acknowledge the realities of the decision making environment and the political nature of evidence. However, it is demonstrated in Chapter 8 that evidence use was described as allowing for symbolic and 'cherry-picked' evidence to be used, rather than rigorous, un-biased evidence incorporated into policy making systematically. This leads back to the thread of discussion throughout the thesis about the need for the more nuanced generation and use of evidence.

It was also found in Chapter 8 that, although an EBPM approach was embodied within regional policy making principles, evaluation was used as a tool for demonstrating accountability, rather than enhancing the effectiveness of RDA policy making through organisational learning. The publication of the IEF did provide a common approach to evaluation (albeit seven years into the operation of the RDAs), however, this chapter demonstrates that the demand for certain types of knowledge led to a central pull for evaluation that embodied a managerialist and mechanistic approach to EBPM. Overall, this empirical study of RDA policy processes affirms Sanderson's (2006) theory of 'instrumental bounded rationality' discussed in Chapter 2.

By investigating the generation, communication and utilisation of evidence, contextual factors (barriers and facilitators) were revealed for the uptake and use of evaluation evidence in practice. Barriers identified included: the lack of a regulatory framework and organisational processes leading to the 'pull' and structured use of evaluation evidence; political backing and support from the 'top' to encourage a culture of evaluation; guidance on planning and evaluating knowledge transfer functions; analytical skill and understanding of the potential and limitations of evaluation evidence; and significant issues with the quality, availability and relevance of monitoring and evaluation data. Facilitators identified included: involving users and evaluators early so that evaluation strategies can be established at the beginning of the lifetime of an organisation; using innovative and targeted dissemination methods; championing evaluation 'from the top'; providing opportunities for continuing professional development; and employing the use of knowledge management systems to support a structured dissemination strategy. It could be presumed that clarifying these factors is important for planning well-designed evaluation processes. In

particular, a potential role for the use of knowledge translation tools was identified within the survey.

10.3.2 The role of knowledge translation tools and decision support

This theme relates to the second research aim and to the third research question.

Research question 3: *What are the potential opportunities and barriers to using a knowledge tool to extend an EBM approach to regional policy investment prioritisation?*

This question is much more normative than the first two questions and includes exploration of how to increase the uptake of evidence. The method employed to investigate this final research question was the development of a decision support tool and the use of an online workshop, conducted with an expert stakeholder group involved in RDA regional policy evaluation (Chapter 3 and 9). The online workshop was designed around two decision making scenarios to compare decision making processes without (scenario 1) and with (scenario 2) the use of the decision support tool. This enabled the exploration of the potential opportunities and barriers to the use of a decision support tool to extend an EBM approach to regional policy investment prioritisation.

Nineteen senior policy makers and analysts from three case study organisations participated in the workshop including: former East Midlands Development Agency (EMDA) officers; former One North East (ONE) officers; and central government officers from the Department of Business, Innovation and Skills (BIS); and the Department for Communities and Local Government (CLG). The workshop was designed to collect standardised quantitative and free text qualitative data and detailed responses were given to the questions, with a total word count across the survey and open forums of over nine thousand words. This qualitative data provided rich detail and the quantitative data indicated the underlying.

It was revealed that decision makers expect to be able to undertake a complex synthesis of information cognitively; combining quantitative data alongside experiential knowledge of deliverability, institutional objectives and political ideas and employing the use of personal judgement. When it came to ranking data within an

investment prioritisation decision making scenario, each of the participants took a different approach and assigned different weightings to the discrete categories of information. It could be inferred that this reflects not only that the analytical ability of each of the decision makers varied, but also that they brought wide-ranging background experience and knowledge to the process.

However, it was also found that there are constraints to the cognitive, analytical processing of information which may provide a rationale for the use of decision support tools. Heuristic bias errors were discussed and it was found that the cognitive synthesis of information becomes compromised as there are increases in the amount of information provided (which may be disparate in relevance, quality and comparability) and as the group decision making environment becomes more complex, involving multiple stakeholder preferences.

A decision support tool was constructed and the process reflected upon within this chapter. It was found that extending an EBM approach for investment prioritisation to the regional policy context, once again, revealed key differences across the sectors in terms of generating and using evaluation evidence. Undertaking this analysis through the lens of programming a decision tool revealed practical analytical and data access implications.

It is argued in Chapter 9 that the development of a knowledge translation tool, operationalised through an evaluation and monitoring framework from the start of an organisation's existence, may facilitate the collection of more appropriate, decision-relevant data linked to an underlying programme theory. It was found that this may enable tangible understanding of how data is to be aggregated, highlight any knowledge gaps and facilitate data-linking to other intelligence sources. The inclusion of policy makers early in the research process may also enable the generation of problem-driven evidence and to shape understanding of how such evidence could be used to support decision making. It was found that a knowledge translation tool would need to be flexible and updated to account for changes in context and the emergence of new evidence. It would also need to be programmed, supported and the relationship between researchers and policy makers facilitated by specialists with both sufficient analytical and communication skills.

Chapter 9 demonstrates that the use of a knowledge translation tool dictating policy or providing 'decisions' ignores the complexity and political nature of decision making. However, it could be argued that their construction and use may enable the more appropriate and systematic utilisation of evaluation evidence to subtly shape policy decision making behaviour within the wider process.

10.3.3 Research management insights from conducting the online workshop

This final theme and the findings discussed are not related to a research aim or to a research question per se, but emerged from the experience of undertaking the research. It was identified in Chapter 3 that the context of the research (RDA abolition) created a situation whereby the study population became hard to reach. To overcome such barriers, the development of the online workshop was a highly innovative approach to data collection, generating lessons from the research management process.

It was found that developing and managing an online workshop is a time intensive process. The use of 'Ning', 'Camtasia', 'GoogleDocs' and 'SurveyMonkey' software to develop the workshop was effective and produced a professional end result, but required researcher training. The importance of using gatekeepers to help recruit participants, keeping the total number of participants to a manageable level and using standardised communication to reduce the duplication of effort was revealed.

Given that such a workshop required participants to have a certain level of digital literacy, and that it was fundamentally important for data to be collected without error, the role of technical support and the need to thoroughly pilot test the workshop was emphasised. In this instance, the use of a 'chat function' to provide technical support was rarely used by participants and was time intensive to manage. It is likely that clear documentation and instructions outlining how to use the workshop were important for reducing the need for such support.

A need was revealed to identify issues which may lead to participant attrition. In recruiting participants for the workshop, there was a benefit in capturing all the information required in one contact (i.e. booking in a time for the telephone interview at the same time as gaining informed consent). It was also essential that all of the distinct elements of the workshop (surveys, online videos, etc.) were embedded within

the workshop web-pages and that the workshop flowed sequentially. Consideration was needed to balance the needs of the research (in terms of the workshop content); against the time it took participants to complete the workshop in reality. Finally, having a telephone interview booked in at the end of the workshop was an effective approach to ensure that participants finished the workshop by the stated deadline.

It was found that to replicate more closely a face-to-face focus group and to generate some level of group dialogue, the use of photographs and open forums could have been more embedded as a mandatory part of the workshop. However, the effect this would have in terms of the time needed to complete the workshop and attrition would need careful consideration and this may have led to issues of anonymity and confidentiality. Overall, running the workshop online and 'asynchronously' was effective at enabling busy, senior and geographically spread individuals to participate in the workshop. It was also cost effective.

10.4 Original contributions to knowledge

The study of an extension of EBM to wider social policy is not a new endeavour (Sefton 2000; Sefton et al. 2002; Sefton 2003; Dobrow, Goel and Upshur 2004; Cookson 2005; Somekh et al. 2005). However, this study makes a first original contribution to knowledge by analysing a case study of EBRPM to explore the wider EBPM debates. Overall a key finding of the study is the need for a more nuanced approach to the generation and use of evidence. This is in contrast to imposing quality criteria specific to one type of study design (e.g. experimental methods) (Chapter 6) and allowing for cherry-picked and unsystematic evidence use within policy making processes (Chapter 8).

This aligns with the call for the appropriate use of evidence identified within health policy (Abeyasinghe and Parkhurst 2013; Rutter, Hawkins and Parkhurst 2013; Nutley, Powell and Davies 2013). However, this thesis makes a second original contribution to knowledge by examining the potential role of knowledge translation tools and decision support in developing an EBPM approach. It is argued that the development of a knowledge translation tool, which is operationalised through an evaluation and monitoring framework from the start of an organisation's existence, should facilitate the collection of more appropriate, decision-relevant data linked to an underlying

programme theory (Chapter 9). It could be argued that this would enable tangible understanding of how data is to be aggregated, highlight any knowledge gaps and facilitate data-linking to other intelligence sources. The inclusion of policy makers early in the research process may also enable the generation of problem-driven evidence and to shape understanding of how such evidence could be used to support decision making.

Finally, due to the context of RDA abolition, the research methods needed to be adapted as the study population became hard to reach (Chapter 3). This leads on to the third original contribution to knowledge. Innovative research methods, in the form of an online workshop, were developed to overcome these barriers, generating original research management insights on the use of such methods. Such a workshop, based upon 'eLearning' approaches went beyond examples of asynchronous online focus groups found in the online research methods literature (e.g. Adriaenssens and Cadman 1999; Bates et al. 2009).

10.5 Limitations of this research

The methodology, and details of why the research was planned and carried out the way it was are detailed in Chapter 3. To assess the overall reliability and validity of this research, triangulation methods have been employed through the use of a mixed-methods design to compare findings across and between the research phases. In terms of reliability, the research findings are internally consistent in that the quantitative and qualitative data collected correspond across the survey and workshop (Chapters 8 and 9) and with the comparative review of academic and policy literature (Chapters 4-6). In terms of external validity, as expected, the results of the research are unlikely to be generalisable. However, the research context and any assumptions have been described to enhance the 'transferability' of the research findings to other settings (Lincoln and Guba 1985).

Overall, the research was undertaken at a time when there was a rapidly evolving policy landscape. Therefore, I do not believe it would be possible to conduct the research again and produce the exact same findings. The qualitative findings rely heavily on people's words and quotations to illuminate a point, and if the study was conducted again, different issues would be of concern to the participants. It might not

even be possible to contact the same individuals. A key point to be made is that this research is heavily focused towards the perspective of the regional agencies, and for the survey towards Yorkshire Forward in particular. If the sampling framework had been biased towards central government officers or other RDAs, different issues may have been raised. Having said that, the 'bottom line conclusions' described in section 10.4, are 'cross-cutting' across both the survey and the workshop, despite Yorkshire Forward officers not participating in the online workshop.

Clearly, the timing of RDA abolition created huge challenges and limitations to the research. The study population became hard to reach and a pragmatic approach needed to be taken and less conventional methodologies applied to meet the aims of the research. The timing of the research is also likely to have had an impact on the participants' responses. Indeed, at the time of the empirical research: RDA officers were facing redundancy, or had already been made redundant; central government officers were facing uncertainty over their future roles; and consultancy firms were facing public sector cuts in expenditure on their services leading to redundancies and a contraction of the sector. There was a sentiment in the collated data that the RDA abolition process would lead to a long-term loss of competence in economic planning and evaluation activities. I inferred that participants felt a range of emotion at this time: uncertainty, loss, anger, defensiveness and also apathy. Although this was taken into consideration when analysing and interpreting the data, it is inevitable that the context of RDA abolition would have impacted upon the participant's responses. In particular, it may have perhaps intensified views about the problems associated with the politicisation of policy processes given the focus in policy discourse on policy termination at the time.

However, undertaking the research at the end of the RDAs' existence also provided considerable opportunities for this research. Individuals that did participate gave detailed and thoughtful responses. I sensed that participants were able to reflect back and critically analyse the RDA 'experiment' as a whole and within the wider policy context, rather than being caught up in the 'minutiae' of day to day working life within an ongoing organisation.

10.6 Implications for policy and practice

This thesis has identified that there have been influential calls for the adoption of a rigorous, evidence based approach to wider social policy making and practice, deploying notions of scientific rigour borrowed from the natural sciences and EBM. In England, the concept of EBPM has been mobilised symbolically by the New Labour, Coalition and Conservative governments through the ‘What Works’ agenda, with evidence portrayed in political discourse as apolitical, neutral and objective. However, the review of the policy and academic literature undertaken has demonstrated that the nature of evidence itself that goes to make up evidence based policy is subject to careful selection influenced by political ideas, institutions and contexts (Chapter 2 and 6). Furthermore, the empirical research undertaken affirms Sanderson’s (2006) theory that instrumental attempts to embed evidence in policy making institutions and practices may lead to ‘instrumental bounded rationality’ and a managerialist and mechanistic approach to policy making (Chapter 8). This is of particular pertinence given the recent election of the Conservative government in May 2015, which signifies a continued dedication to austerity reforms and the likely further squeeze on research and evaluation budgets. Given the commitment made to EBM approaches and the What Works Centres during the Coalition’s administration, it can be presumed that economic and social policy evaluation practice is likely to face further pressure to focus more narrowly on RCTs, economic evaluation and evidence synthesis.

When examining the relevance of this for evaluation practice with respect to the current arrangements for sub-national economic development in England, it is apposite to reflect upon the rise of ‘new spatial economics’ (Martin et al. 2015) and the challenge towards place-based policies (most notably by Overman and colleagues at the WWG). More importantly, however, is that further cuts in government spending will likely mean a shift away from people focused interventions too. Thus the evaluation of regional policy is at somewhat of a crossroads and faces an uncertain future. Yet, when reflecting upon the recently published BIS Evaluation Strategy 2015-16 (BIS 2014b), it can be presumed that the issues raised in Chapters 5, 8 and 9 over the quality and consistency of regional policy evaluation outputs will be compounded as the policy landscape becomes more complex, moving from the evaluation of 9 RDAs compared to 39 LEPS.

Probably one of the most interesting aspects of this research is that links to the work programme for the WWG are evident. The empirical research conducted on decision support was underpinned by developing a model where (theoretically) interventions could be sorted by cost-effectiveness ratios with a threshold applied at the point when the RDA budget is exhausted. This was based upon a simplified version of NICE decision making processes (see Chapter 4). The empirical analysis reported upon in Chapter 9 draws upon the views of policy makers and analysts to critically analyse such an approach. This work was undertaken prior to the announcement of the 'What Works Centres' conceptualised collectively as a "NICE for Social Policy" (Cabinet Office 2013, p1). Through developing the decision support tool (Chapter 9), it was found that the task of synthesising the evidence base and programming a decision logic into the tool to enable users to compare policies according to a 'common currency' (i.e. a common set of standards in each area for comparing the effectiveness of interventions) is currently unfeasible. Considerable methodological, theoretical and practical constraints were identified through the empirical analyses presented in Chapters 6, 8 and 9. In turn, this has a direct implication for the WWG's mandate to develop an online toolkit as part of its knowledge transfer and exchange work programme.

10.7 Recommendations for future research

There are a number of potential avenues for future research:

- Although regional/local policy has subsets and related fields, such as cluster policy, small business policy, skills policy and infrastructure planning, and a number of subnational institutions deliver economic policies (each of which may be analysed using quasi-experimental approaches); this research has focused on regional policy as delivered through the RDAs. Future research could widen out the analysis beyond the scope of this thesis.
- Analysis of the role of decision support was conducted through the online workshop. However, to test the validity and relevance of such an approach, future research could investigate the role of a decision support tool linked to an ongoing evaluation framework 'in the field' to explore the influence on the utilisation and impact of evidence, as well as any negative effects.

- Learning from considering the context in regional development could also be helpful in feeding back into the health context as it changes and broadens to extend the EBM approach to wider health and social care policy. Future research could widen out the analysis to look at potential opportunities for how the health sector, and NICE in particular, could make use of such findings.
- The online workshop method certainly merits further application, especially in situations involving 'hard to reach' study populations involving senior, busy and geographically spread individuals.
- Eighteen telephone interviews were conducted and direct feedback was given by participants on the use of the online workshop as part of this PhD. The interviews were recorded, but the data was not transcribed. Given the volume and quality of data collected from the online survey and workshop, the interview data has not been included as part of this thesis. Such data could, however, form the basis of future research work.

10.8 Concluding remarks

To conclude this thesis, and reflect upon the (appropriate) generation of evidence, it is apposite to reconsider Hugo's (1862, p1004) quotation from the start of Chapter 1:

Have courage, citizens! We must go forward. But what are we aiming at? At government by knowledge, with the nature of things the only social force...

To interpret society 'going forward' through the lens of policy evaluation, building upon the insights from this study, one could advocate the need to invest in a longitudinal evidence base, build upon institutional memory and undertake innovative work on evaluation and decision support methodology to improve the quality and relevance of evidence. However, given the familiar cycle of political and institutional churn embodied in UK politics, such a long-term approach is unlikely. Instead, it has been demonstrated in this thesis that the 'what works' agenda has tended to embody a managerialist manifestation of EBPM, which has focused on the mechanistic supply of evidence, rather than on a coordinated learning process to support an enabling culture of evidence informing policy decisions.

Bibliography

ABEYSINGHE, S. and PARKHURST, J. (2013). *Better evidence for policy: from hierarchies to appropriateness*, GRIP-Health Working Papers. London, London School of Hygiene and Tropical Medicine.

ADRIAENSSENS C. and CADMAN L. (1999). An adaptation of moderated e-mail focus groups to assess the potential for a new online (Internet) financial services offer in the UK. *Journal of the Market Research Society*, **41** (4), 417-424.

ALLEN, T., GRACE, C. and MARTIN, S. (2014). *The role of evidence and research in local government*. Local Government Knowledge Navigator. [online]. Last accessed 5 January 2015 at: http://www.solace.org.uk/knowledge/reports_guides/LGKN-Research-survey-FINAL.pdf

ALLIANCE FOR USEFUL EVIDENCE (2014). *The NICE way: lessons for social policy and practice from the national institute for health and care excellence*. [online]. Last accessed 4 December 2014 at: https://www.nice.org.uk/Media/Default/News/NestaAlliance_and_NICE_paper.pdf

ALLSOP, J. (1995). *Health Policy and the NHS: towards 2000*. 2nd ed., Harlow, Longman.

ARMSTRONG, H. and TAYLOR, J. (2000). *Regional Economics and Policy*. 3rd ed., Oxford, Blackwell.

ARMSTRONG, H. and WELLS, P. (2006a). Evaluating the governance of structural funds programmes: The case of community economic development in South Yorkshire. *European Planning Studies*, **14**(6), 855-876.

ARMSTRONG, H. and WELLS, P. (2006b). Structural Funds and the evaluation of community economic development initiatives in the UK: A critical perspective. *Regional studies*, **40**, 259-272.

ARMSTRONG, H., KEHRER, B., WELLS, P. and WOOD, A. (2002). The evaluation of community economic development initiatives. *Urban Studies*, **39**, 457-481.

ARMSTRONG, R., WATERS, E., CROCKETT, B and KELEHER, H. (2007) The nature of evidence resources and knowledge translation for health promotion practitioners. *Health Promotion International*, **22** (3), 254-260.

ARONSON, E. and CARLSMITH, J. (1968). Experimentation in social psychology. In: LINDZEY, G. and ARONSON, E. (eds.). *The handbook of social psychology*, vol.2. Reading MA, Addison-Wesley, 1-79.

ATKINS, D., ECCLES, M., FLOTTORP, S., GUYATT, G., HENRY, D., HILL, S., LIBERATI, A., O'CONNELL, D., OXMAN, A., PHILLIPS, B., SCHÜNEMANN, H., EDEJER, T., VIST, G. and WILLIAMS, J. (2004). Systems for grading the quality of evidence and the strength

of recommendations 1: critical appraisal of existing approaches. The Grade Working Group. *BMC Health Services Research*, **4** (1), 38.

AUDRETSCH, D. and FELDMAN, M. (1996). Innovative clusters and the industry life cycle. *Review of Industrial Organisation*, **11**, 253-273.

BACHE, I. (1998). *The Politics of European Union Regional Policy: Multi-Level Governance or Flexible Gatekeeping?* Contemporary European Studies Series, Sheffield, UACES/Sheffield Academic Press.

BACHTLER, J. (2001). Where is regional policy going? *Changing concepts of regional policy. EPRC discussion paper*, University of Strathclyde. [online]. Last accessed 3 May 2014 at: <http://www.eprc.strath.ac.uk/eorpa/reports2.cfm>

BACHTLER, J. (2011). The Evaluation of Regional Policy in Europe: Culture, Commitment and Capacity. In: BIENIAS S., OLEJNICZAK K. and KOZAK, M. (eds.). *Evaluating the Effects of Regional Interventions*. Warsaw, Ministry of Regional Development, 94-118.

BACHTLER, J. and MICHIE, R. (1995). A New Era in EU Regional Policy Evaluation? The Appraisal of the Structural Funds. *Regional Studies*, **29** (8), 745-751.

BACHTLER, J. and WREN, C. (2006). Evaluation of European Union Cohesion Policy: Research Questions and Policy Challenges. *Regional Studies*, **40** (2), 143-154.

BACHTLER, J., MENDEZ, C. and VIRONEN, H. (2014). *Regional development and policy in Europe*, Collection Studies No.2. Madrid, EUROsocial.

BAGGOTT, R. (1998). *Health and Health Care in Britain*. Basingstoke, Macmillan.

BAGSHAW, S. and BELLOMO, R. (2008). The need to reform our assessment of evidence from clinical trials: A commentary. *Philosophy, Ethics, and Humanities in Medicine*, **3**, 23.

BAGSHAW, S. and BELLOMO, R. (2008). The need to reform our assessment of evidence from clinical trials: A commentary. *Philosophy, Ethics, and Humanities in Medicine*, **3**, 23.

BARCA, F. (2009). *An Agenda for a Reformed Cohesion Policy, A Place-Based Approach to Meeting European Union Challenges and Expectations*, Independent report to the European Commission. Brussels, DG Regio.

BARCA, F. (2011). *OECD Regional Outlook 2011: Building Resilient Regions for Stronger Recovery*. Paris, Organisation for Economic Cooperation and Development.

BARLOW REPORT (1940). *Royal Commission on the Distribution of the Industrial Population Report*. London, HMSO.

BARNES, A. and PARKHURST, J. (2014). Can global health policy be depoliticised? A critique of global calls for evidence-based policy. In: (eds.) BROWN, G., YAMEY, G. and WAMALA, S. (eds.). *The Handbook of Global Health Policy*. Cambridge, Wiley-Blackwell.

BARRO, R. and SALA-I-MARTIN, X. (1991). Convergence across states and regions. *Brookings Papers on Economic Activity*, **1**, 107-182.

BAUMBUSCH, J., KIRKHAM, S., KHAN, K., MCDONALD, H., SEMENIUK, P., TAN, E. and ANDERSON, J. (2008). Pursuing common agendas: A collaborative model for knowledge translation between research and practice in clinical settings. *Research in Nursing & Health*. **31**(2), 130-140.

BAZERMAN, M. (1998). *Judgment in Managerial Decision Making*. 4th ed., New York, John Wiley and Sons.

BEATTY, C., FOTHERGILL, S. and GORE, T. (2012) *The real level of unemployment 2012*. Sheffield, CRESR, Sheffield Hallam University [online]. Last accessed 7 April 2015 at: <http://www.shu.ac.uk/research/cresr/sites/shu.ac.uk/files/real-level-of-unemployment-2012.pdf>

BEGG, I., BACHTLER, J., MACCHIARELLI, C., MENDEZ, C. and WISHLADE, F. (2014). *European Economic Governance and Cohesion Policy*, Report to the European Parliament Committee on Regional Development [online]. Last accessed 7 April 2015 at: http://www.europarl.europa.eu/RegData/etudes/etudes/join/2014/513999/IPOL-REGI_ET%282014%29513999_EN.pdf

BELSKY, J., BARNES, J. and MELHUISH, E. (eds.) (2007). *The national evaluation of Sure Start: Does area-based early intervention work?* Bristol, Policy Press.

BENNEWORTH, P. (2011) Five Lessons to Learn From the English Regional Development Agencies. *Regions Magazine*, **284** (1) 11-13.

BENTLEY, G., BAILEY, D. and SHUTT, J. (2010). From RDAs to LEPs: A New Localism? Case Examples of West Midlands and Yorkshire. *Local Economy*, **25** (7), 535-557.

BERELSON, B. (1952). *Content Analysis in Communication Research*. New York, Free Press.

BLAGESCU, M., DE LAS. CASAS, L. and LLOYD, R. (2005). *Pathways to Accountability: The GAP Framework*. One World Trust, London [online]. Last accessed 7 April 2015 at: <http://www.who.int/management/partnerships/accountability/PathwaysAccountabilityGAPFramework.pdf>

BORENSTEIN, M., HEDGES, L., HIGGINS, J. and ROTHSTEIN, H. (2009). *Introduction to meta-analysis*. Chichester UK, Wiley.

- BOVAIRD, T. (2014). Attributing Outcomes to Social Policy Interventions – ‘Gold Standard’ or ‘Fool’s Gold’ in Public Policy and Management? *Social Policy & Administration*, **48** (1), 1-23.
- BRIGGS, A. (2000). Handling uncertainty in cost-effectiveness models. *Pharmacoeconomics*, **22**, 479-500.
- BRIGGS, A., CLAXTON, K. and SCULPHER, M. (2006). *Decision modelling for health economic evaluation*. Oxford, Oxford University Press.
- BRIGHT, T., WONG, A., DHURJATI, R., BRISTOW, E., BASTIAN, L., COEYTAUX, R., SAMSA, G., HASSELBLAD, V., WILLIAMS, J., MUSTY, M., WING, L., KENDRICK, A., SANDERS, G. and LOBACH, D. (2012). Effect of clinical decision-support systems: a systematic review. *Annals of Internal Medicine*, **157**, 29-43.
- BRYAN, S., WILLIAMS, I. and MCLVER, S. (2007). Seeing the NICE side of cost-effectiveness analysis: a qualitative investigation of the use of CEA in NICE technology appraisals. *Health Economics*, **16** (12), 179-193.
- BRYMAN, A. (2012). *Social Research Methods*. 2nd ed., Oxford, Oxford University Press.
- BUCHKOWSKY, S. and JEWESSON, P. (2004). Industry sponsorship and authorship of clinical trials over 20 years. *The Annals of Pharmacotherapy*, **38**, 579-85.
- CABINET OFFICE (1999a). *Modernising Government*. CM4310. London, TSO.
- CABINET OFFICE (1999b). *Professional Policy Making for the Twenty-First Century*. London, Cabinet Office Strategic Policy Making Team.
- CABINET OFFICE (1999c). *Sharing the Nation’s Prosperity: Variations in Economic and Social Conditions across the UK*. London, Cabinet Office.
- CABINET OFFICE (2011). *Open Public Services White Paper*. London, Cabinet Office.
- CABINET OFFICE (2013). *What Works: Evidence Centres for Social Policy*. London, TSO.
- CANTER, D. (2012). *Beyond the Randomised Controlled Trial*. [online]. Last accessed 9 December 2014 at: <http://www.socialsciencespace.com/2012/08/beyond-the-randomised-controlled-trial/>
- CARMINES, E. and ZELLER, R. (1991). *Reliability and Validity Assessment*. Newbury Park CA, Sage.
- CARTWRIGHT, N. and HARDIE, J. (2012). *Evidence-based policy: A practical guide to doing it better*. Oxford, Oxford University Press.
- CENTRE FOR REVIEWS AND DISSEMINATION [CRD]. (2008). *Systematic Reviews: CRD’s guidance for undertaking reviews in healthcare*. [online]. Last accessed 23 October 2014 at: http://www.york.ac.uk/inst/crd/pdf/Systematic_Reviews.pdf

CHADWICK, M. (2014). Former RDA National Secretariat. *RDA evaluation*. Interview with the author, 14th October 2014. Personal communication.

CHADWICK, M., TYLER, P. and WARNOCK, C. (2013). How to raise the bar on impact evaluation: Challenges for the evaluation of local enterprise partnerships and the regional growth fund in times of austerity. *Local Economy*, **28** (7-8), 842-85.

CHALMERS, I. (2007). The lethal consequences of failing to make full use of all relevant evidence about the effects of medical treatments: the importance of systematic reviews. In: ROTHWELL P. (ed). *Treating individuals: from randomised trials to personalized medicine*. London, The Lancet.

CLARK, R. (2011). Senior Librarian: eLearning and eResources at Leeds College of Music. *eLearning*. Interview with the author, 23rd November 2011. Personal communication.

CLARKE, S. (1988). *Keynesianism, Monetarism and the Crisis of the State*. London, Edward Elgar.

CLAXTON, K., SCULPHER, M., MCCABE, C., BRIGGS, A., BUXTON, M., BRAZIER, J., AKEHURST, R. and O'HAGAN, A. (2005). Probabilistic sensitivity analysis for NICE technology assessment: not an optional extra. *Health Economics*, **14**, 339-347.

CLAY, E. and SCHAFFER, B. (1984). *Room for Manoeuvre: An Exploration of Public Policy in Agricultural and Rural Development*. London, Heineman Educational Books.

COCHRANE, A. (1972). *Effectiveness and Efficiency: Random Reflections on Health Services*. [online]. Last accessed 23 October 2014
at: http://www.nuffieldtrust.org.uk/sites/files/nuffield/publication/Effectiveness_and_Efficiency.pdf

COCHRANE, A. (1979). 1931-1971: a critical review, with particular reference to the medical profession. In: TEELING-SMITH, G. and WELLS, N. (eds.). *Medicines for the year 2000*. London, Office of Health Economics, 1-11.

COHEN, M. MARCH, J. and OLSEN, J. (1972). A garbage can model of organizational choice. *Administrative Science Quarterly*, **17** (1), 1-25.

COLBY, D., QUINN, B., WILLIAMS, C., BILHEIMER, L. and GOODELL, S. (2008). Research Glut and Information Famine: Making Research Evidence More Useful for Policymakers. *Health Affairs*, **27** (4), 1177-1182.

COLLIER, D., (1993). The Comparative Method. In: FINIFTER, A. (ed.) *Political Science: The State of Discipline II*. Washington, DC: American Political Science Association, 105-119.

COLLIER, P. and DOLLAR, D. (2001). *Globalization, Growth and Poverty: Building an Inclusive World Economy*. Oxford, Oxford University Press.

- CONNELL, J. and KUBISCH, A. (1998). Applying a Theory of Change Approach. In: FULBRIGHT-ANDERSON, K., KUBISCH, A. and CONNELL, J. (eds.). *New Approaches to Evaluating Community Initiatives Volume 2: Theory, Measurement, and Analysis*. Washington DC, The Aspen Institute.
- CONTANDRIOPOULOS, D., LEMIRE, M., DENIS, J.L. and TREMBLAY, E. (2010). Organizations and Policy Arenas: A Narrative Systematic Review of the Literature. *The Milbank Quarterly*, **88**, 444-483.
- COOK, J., DICKINSON, S., HEUMAN, D. and WHITE, G. (2008). *Viewpoint: Making the most of evaluation*, Issue 3, SQW Group. [online]. Last accessed 3 May 2014 at: [http://www.sqw.co.uk/files/2713/8715/0441/Viewpoint - Making the most of Evaluation.pdf](http://www.sqw.co.uk/files/2713/8715/0441/Viewpoint_-_Making_the_most_of_Evaluation.pdf)
- COOKSON, R. (2005). Evidence-based policy making in health care: what it is and what it isn't. *Journal of Health Services Research & Policy*, **10**, 118-121.
- COURT, J. and YOUNG, J. (2003) Bridging research and policy: Insights from 50 case studies. [online] Last accessed 07 April 2015 at: http://ewc-popcomm.org/pdf/2_Selected_Readings/Bridging_research_policy_ODI.pdf
- CRAFTS, N. (2005). Regional GDP in Britain 1871- 1911: some estimates. *Scottish Journal of Political Economy*, **52** (1), 54-64.
- CRESWELL, J. and CLARK, P. (2011). *Designing and conducting mixed methods*. 2nd ed., Thousand Oaks CA, Sage.
- CULYER, A. and NEWHOUSE, J. (eds.). (2000). *Handbook of Health Economics*, vol. 1A. Amsterdam, Elsevier.
- CULYER, A., MCCABE, C., BRIGGS, A., CLAXTON, K., BUXTON, M., AKEHURST, R., SCULPHER, M. and BRAZIER, J. (2007). Searching for a threshold, not setting one: the role of the National Institute of Health and Clinical Excellence. *Journal of Health Services Research and Policy*, **12** (1), 56-58.
- DAVIES, H., NUTLEY, S. and SMITH, P. (eds.) (2000). *What works? Evidence-based policy and practice in public services*. Bristol, Policy Press.
- DAVIES, P. (2004). *Is Evidence-Based Government Possible?* Jerry Lee Lecture presented to the 4th Annual Campbell Collaboration Colloquium. [lecture]. Held on 19 February, Washington USA.
- DAVIS, P. (2005). The Limits of Realist Evaluation: Surfacing and, Exploring Assumptions in Assessing the Best Value Performance Regime, *Evaluation*, **11** (3), 275-295.
- DAVISON, C. (2013). *Critical Examination of Knowledge to action models and implications for promoting Health Equity*. Antigonish, National Collaborating Centre for Determinants of Health.

- DEPARTMENT FOR BUSINESS, INNOVATION AND SKILLS [BIS]. (2009a). *RDA evaluation: Practical Guidance on Implementing the Impact Evaluation Framework*. London, BIS.
- DEPARTMENT FOR BUSINESS, INNOVATION AND SKILLS [BIS]. (2009b). *RDA evaluation: Practical Guidance on Implementing the Impact Evaluation Framework. Appendix 1 – Beneficiary survey methodology and questionnaires*. London, BIS.
- DEPARTMENT FOR BUSINESS, INNOVATION AND SKILLS [BIS]. (2009c). *Research to Improve the Assessment of Additionality*, BIS, Occasional Paper No. 1. Cambridge Economic Associates. London, BIS.
- DEPARTMENT FOR BUSINESS, INNOVATION AND SKILLS [BIS]. (2009d). *Guidance for Using Additionality Benchmarks in Appraisal*. London, BIS
- DEPARTMENT FOR BUSINESS INNOVATION AND SKILLS [BIS]. (2010a). *Local Growth: Realising Every Place's Potential*. London, BIS.
- DEPARTMENT FOR BUSINESS INNOVATION AND SKILLS [BIS] (2010b). *Understanding Local Growth*, BIS economics paper no 7. London, BIS.
- DEPARTMENT FOR BUSINESS, INNOVATION AND SKILLS [BIS] (2011). *Innovation and Research Strategy*. London, TSO.
- DEPARTMENT FOR BUSINESS, INNOVATION AND SKILLS [BIS] (2013). *Local Enterprise Partnerships*. [online]. Last accessed 23 October 2014 at:
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/252793/bis-11-768-local-enterprise-partnerships-boundary-map-august-2013.pdf
- DEPARTMENT FOR BUSINESS, INNOVATION AND SKILLS [BIS] (2014a). *Guidance, Regional Growth Fund*. [online]. Last accessed 23 October 2014 at:
<https://www.gov.uk/understanding-the-regional-growth-fund>
- DEPARTMENT FOR BUSINESS, INNOVATION AND SKILLS [BIS] (2014b). *Evaluation Strategy 2015-16: Annexes, Accountability and learning at the heart of BIS*. [online]. Last accessed 28 January 2015 at:
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/387512/bis-14-1296-evaluation-strategy-2015-16-annexes-accountability-and-learning-at-the-heart-of-bis.pdf
- DEPARTMENT FOR COMMUNITIES AND LOCAL GOVERNMENT [CLG] (2010a). *The New Deal for Communities Experience, a final assessment*. London, CLG.
- DEPARTMENT FOR COMMUNITIES AND LOCAL GOVERNMENT [CLG] (2010b). *Evaluation of the National Strategy for Neighbourhood Renewal*. London, CLG.
- DEPARTMENT OF HEALTH (1996). *Promoting Clinical Effectiveness: A Framework for Action in and through the NHS*. London, Department of Health.

DEPARTMENT OF TRADE AND INDUSTRY [DTI] (2003). *Single Programme Appraisal Guidance*. Version 7f. London, TSO.

DEPARTMENT OF TRADE AND INDUSTRY [DTI] (2006). *Evaluating the impact of England's regional development agencies: Developing a methodology and evaluation framework*, DTI Occasional Paper No. 2. London, Department of Trade and Industry.

DIEZ, Maria-Angeles (2002). Evaluating New Regional Policies. Reviewing the Theory and Practice. *Evaluation*, **8** (3), 285-305.

DILLMAN, D. (2007). *Mail and Internet surveys - The tailored design method*. 2nd ed., New York, John Wiley & Sons.

DIRECTORATE-GENERAL FOR REGIONAL AND URBAN POLICY [DG REGIO]. (2014a). *What is regional policy? History of the policy*. [online]. Last accessed 23 October 2014 at: http://ec.europa.eu/regional_policy/what/milestones/index_en.cfm

DIRECTORATE-GENERAL FOR REGIONAL AND URBAN POLICY [DG REGIO]. (2014b). *Evaluation: History of the policy*. [online]. Last accessed 23 October 2014 at: http://ec.europa.eu/regional_policy/impact/evaluation/library_en.cfm

DIXON-WOODS, M., AGARWAL, S., JONES, D., YOUNG, B. and SUTTON, A. (2005) Synthesising qualitative and quantitative evidence: a review of possible methods. *Journal of Health Services Research and Policy*, **10** (1), 45-53.

DOBROW, M. J., GOEL, V. and UPSHUR, R. (2004). Evidence-based health policy: context and utilisation. *Social Science and Medicine*, **58**, 207-218.

DODGSON, J., SPACKMAN, M., PEARMAN, A. and PHILLIPS, L. (2009). *Multi-criteria analysis: a manual*. London, Department for Communities and Local Government.

DRIESSEN, S., HUIJSEN, W. and GROOTVELD, M. (2007). A framework for evaluating knowledge- mapping tools. *Journal of Knowledge Management*, **11** (2), 109-117.

DRUMMOND, M., SCULPHER, M., TORRANCE, G., O'BRIEN, B. and STODDART, G. (2005). *Methods for the Economic Evaluation of Health Care Programmes*. 3rd ed., Oxford, Oxford University Press.

EAST MIDLANDS DEVELOPMENT AGENCY [EMDA] (2012). *Annual Report and Accounts 2011 – 2012*. [online]. Last accessed 5 December 2014 at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/247065/0174.pdf

EDWARDS, P., ROBERTS, I., CLARKE, M., DIGUISEPPI, M., PRATAP, S., WENTZ, R. and KWAN, I. (2002). Increasing response rates to postal questionnaires: systematic review. *British Medical Journal*, **324**, 1183-1191.

ENGEL, G. (1977). The need for a new medical model: A challenge for biomedicine. *Science*, **196**, 129-136.

ENGLISH PARTNERSHIPS (2008). *A Standard Approach to Assessing the Additional Impact of Projects Method Statement*, 3rd ed. London, English Partnerships.

EOM, S. and KIM, E. (2006). A survey of decision support system applications (1995–2001). *Journal of the Operational Research Society*, **57**, 1264-1278.

ESTABROOKS, C., DAVID, S., LOVELY, J., JACQUE, E. and HOFMEYER, A. (2006). A Guide to knowledge translation theory. *The Journal of Continuing Education in the Health Professions*, **26**, 25-36.

EUROPEAN COMMISSION (1995). *MEANS Handbook No.3: Measuring the employment effects of community structural interventions*. Luxembourg, Office for Official Publications of the European Communities.

EUROPEAN COMMISSION (1999). *The MEANS Collection: Evaluating Socio-Economic Programmes*. Luxembourg, Office for Official Publications of the European Communities.

EUROPEAN COMMISSION (2006a). *Indicative Guidelines on Evaluation Methods: Ex-Ante Evaluation*, Working Document No 1. Directorate General for Regional Policy. [online]. Last accessed 3 May 2014 at:
http://ec.europa.eu/regional_policy/information/evaluations/guidance_en.cfm#3

EUROPEAN COMMISSION (2006b). *Guidance on the methodology for carrying out cost-benefit analysis*, Working Document No 4. Directorate General for Regional Policy. [online]. Last accessed 3 May 2014 at:
http://ec.europa.eu/regional_policy/index_en.htm

EUROPEAN COMMISSION (2007). *Indicative Guidelines on Evaluation Methods: Evaluation during the Programming Period*, Working Document No 5. Directorate General for Regional Policy. [online]. Last accessed 3 May 2014 at: 2014 at:
http://ec.europa.eu/regional_policy/information/evaluations/guidance_en.cfm#3

EUROPEAN COMMISSION (2013). *EVALSED: The resource for the evaluation of Socio-Economic Development*. [online]. Last accessed 1 October 2014 at:
http://ec.europa.eu/regional_policy/sources/docgener/evaluation/guide/guide_evalsed.pdf

EUROQOL (2014). *What is EQ-5D?* [online]. Last accessed 23 October 2014 at:
<http://www.euroqol.org/>

EVANS, D. (2003). Hierarchy of evidence: A framework for ranking evidence evaluating health care interventions. *Journal of Clinical Nursing*, **12**, 77-84.

EVERY-PALMER, S. and HOWICK, J. (2014). How evidence-based medicine is failing due to biased trials and selective publication. *Journal of Evaluation in Clinical Practice*. [online], in press, 12 May 2014. Article from Wiley and Sons last accessed 4 Dec 2014 at: <http://onlinelibrary.wiley.com/doi/10.1111/jep.12147/abstract>

- EVIDENCE-BASED MEDICINE WORKING GROUP [EBM WORKING GROUP]. (1992). Evidence-based medicine: A new approach to teaching the practice of medicine. *JAMA*, **268**, 2420-2425.
- FENWICK, E., CLAXTON, K. and SCULPHER, M. (2001). Representing uncertainty: the role of cost-effectiveness acceptability curves. *Health Economics*, **10**, 779-787.
- FERLIE, E. and MCGIVERN, G. (2014). Bringing Anglo-Governmentality into Public Management Scholarship: The Case of Evidence-based Medicine in UK Health Care. *Journal of Public Administration: Research and Theory*, **24**, 59-83.
- FERLIE, E., FITZGERALD, L., MCGIVERN, G., DOPSON, S. and BENNETT, C. (2013). *Making Wicked Problems Governable? The Case of Managed Networks in Health Care*. Oxford, Oxford University Press.
- FERRY, M. and BACHTLER J. (2013). Reassessing the concept of policy termination: The case of regional policy in England. *Policy Studies*, **34** (3), 255-273.
- FESER, E. (2013). Isserman's Impact: Quasi-Experimental Comparison Group Designs in Regional Research. *International Regional Science Review*, **36**(1), 44-68.
- FOLLAND, S., GOODMAN, A. and STANO, M. (2010). *The Economics of Health and Health Care*. 6th ed., Boston, Prentice Hall.
- FOOTE, C., GOETTE, L. and MEIER, S. (2009). *Policymaking Insights from Behavioural Economics*. Boston, Federal Reserve Bank of Boston.
- FRENCH, S., MAULE, J. and PAPAMICHAIL, K. (2009). *Decision Behaviour, Analysis and Support*. Cambridge, Cambridge University Press.
- GARDINER, B., MARTIN, R. and TYLER, P. (2004). Competitiveness, productivity and economic growth across the European Regions. *Regional studies*, **38** (9), 1045-1067.
- GARDINER, B., MARTIN, R. and TYLER, P. (2012). *Spatially Unbalanced Growth in the British Economy*, CGER Working Paper Number 1. Cambridge, University of Cambridge Centre for Geographical Economic Research (CGER).
- GARRETSEN, H., MCCANN, P., MARTIN, R. and TYLER, P. (2013). The future of regional policy. *Cambridge Journal of Regions, Economy and Society*, **6**, 179-186.
- GERBER, A., HENTZELT, F., and LAUTERBACH, K. (2007). Can evidence-based medicine implicitly rely on current concepts of disease or does it have to develop its own definition? *Journal of Medical Ethics*, **33**(7), 394-99.
- GILLESPIE G., MCGREGOR P., SWALES J. and YIN Y. (2001). The displacement and multiplier effects of Regional Selective Assistance: a computable general equilibrium analysis. *Regional Studies*, **35**, 125-139.

- GIVEN, L. (ed.) (2008). *The SAGE encyclopaedia of qualitative research methods*. Thousand Oaks CA, Sage.
- GLAESER, E. and RESSEGER, M. (2010). The complementarity between cities and skills. *Journal of Regional Science*, **50** (1), 221-244.
- GLASBY, J. and BERESFORD, P. (2006). Who knows best? Evidence-based practice and the service user contribution. *Critical Social Policy*, **26** (1), 268-284.
- GLASS, G. (1976). Primary, secondary and meta-analysis of research. *Educational Researcher*, **5**, 3-8.
- GLENNY, A., ALTMAN, D., SONG, F., SAKAROVITCH, C., DEEKS, J., D'AMICO, R., BRADBURN, M. and EASTWOOD, A. (2005). Indirect comparisons of competing interventions. *Health Technology Assessment*. **9** (26), 1-148.
- GOLDACRE, B. (2013). *Bad Science: Here's our Cabinet Office paper on randomised trials of government policies. Read it.* [online]. Last accessed 9 June 2014 at: <http://www.badsience.net/category/evidence-based-policy/>
- GOODMAN, N. (1999). Who will challenge evidence-based medicine? *Journal of the Royal College of Physicians*, **33**, 249-251.
- GORRY, G. and SCOTT MORTON, M. (1971). A framework for management information systems. *Sloan Management Review*, **13**(1), 1-22.
- GOULD, N. and KENDALL, T. (2007). Developing the NICE/SCIE Guidelines for Dementia Care: The Challenges of Enhancing the Evidence Base for Social Care. *British Journal of Social Work*, **37**, 475-490.
- GRAHAM, I., LOGAN, J., HARRISON, M., STRAUS, S., TETROE, J., CASWELL, W. and ROBINSON, N. (2006). Lost in Knowledge Translation: Time For A Map? *J Contin Educ Health Prof*, **26** (1), 13.
- GRAY, J. (1997). *Evidence-Based Healthcare: How to Make Health Policy and Management Decisions*. London, Churchill Livingstone.
- GREAT BRITAIN (1998). *Regional Development Agencies Act 1998*. [online]. Chapter 45. Office of Public Sector Information. Last Accessed 3 May 2014 at: http://www.opsi.gov.uk/Acts/acts1998/ukpga_19980045_en_1
- GREAT BRITAIN (2012). *Health and social care act 2012*. [online]. Chapter 7. London, The Stationery Office. Last accessed 13 June 2012 at: <http://www.legislation.gov.uk/ukpga/2012/7/contents/enacted>
- GREAT BRITAIN, Parliament, House of Commons (2008). Select Committee on Health: *first report of session 2008-9*. [online]. London, TSO. HC 27-I. Last accessed 4 December 2014 at:

GREAT BRITAIN, Parliament, House of Commons (2010). *Regenerating the English regions: Regional Development Agencies' Support to Physical Regeneration Projects*. London, TSO. HC 214 (2009-10).

GREAT BRITAIN, Parliament, House of Commons (2015). *Local Enterprise Partnerships*. [online]. London, TSO. SN/EP/5651. Last accessed 29 July 2015 at: <http://researchbriefings.files.parliament.uk/documents/SN05651/SN05651.pdf>

GREENE, J. (1994). Qualitative programme evaluation: practice and promise. In: DENZIN, N. and LINCOLN, Y. (eds.). *Handbook of qualitative research*. Thousand Oaks CA, Sage, 530-544.

GREENHALGH, T. and RUSSELL, J. (2006). Reframing evidence synthesis as rhetorical action in the policy making drama. *Healthcare policy*, **1**, 34-42.

GREENHALGH, T., HOWICK, J. and MASKREY N. (2014). Evidence based medicine: a movement in crisis? *BMJ*, **348**, 3725.

GRIMSHAW D. and RUBERY J. (2012). The end of the UK's liberal collectivist social model? The implications of the Coalition government's policy during the austerity crisis. *Cambridge Journal of Economics*, **36** (1), 105-126.

GRIMSHAW, J., ECCLES, M., THOMAS, R., MACLENNAN, G., RAMSAY, C., FRASER, C. and VALE, L. (2006), Toward Evidence-Based Quality Improvement. *Journal of General Internal Medicine*, **21** (S2): S14-S20.

GRIPAIO, P. and BISHOP, P. (2006) Objective One funding in the UK: a critical assessment. *Regional Studies*, **40** (8) 937-952.

GROSSMAN, J. and MACKENZIE, F. (2005). The Randomized Controlled Trial: gold standard, or merely standard? *Perspectives in Biology and Medicine*, **48**, 516-534.

GUBA, E. (1981) Criteria for Assessing the Trustworthiness of Naturalistic Inquiries. *Educational Communication and Technology*, **29**(2), 75-91.

GUYATT, G. (1991). Evidence based medicine. *ACP Journal Club*, (SuppA-16), 114.

GUYATT, G., RENNIE, D., MEADE, M. and COOK, D. (2008). *Users guide to the medical literature: a manual for evidence-based clinical practice*. 2nd ed., US, McGraw Hill and the American Medical Association.

HALLSWORTH, M., PARKER, S. and RUTTER, J. (2011). *Policy making in the real world*. London, Institute for Government.

HARRISON, S. (1998). The politics of evidence-based medicine in the United Kingdom. *Policy & Politics*, **26**, 15-31.

HART, M., DRIFFIELD, N., ROPER, S. and MOLE, K. (2008). *Evaluation of Regional Selective Assistance (RSA) and its successor, Selective Finance for Investment in England (SFIE)*, BERR Occasional Paper No. 2. London, BERR.

HAYNES, L., SERVICE, O., GOLDACRE, B. and TORGERSON, D. (2012). *Test, Learn, Adapt: Developing Public Policy with Randomised Controlled Trials*. London, Cabinet Office and Behavioural Insights Team.

HAYNES, R., DEVEREAUX, P. and GUYATT, G. (2002). Clinical expertise in the era of evidence-based medicine and patient choice. *ACP Journal Club*, **136**, A11-A14

HAYNES, R., SACKETT, D., GRAY, J., COOK, D. and GUYATT, G. (1996). Transferring evidence from research into practice. 1. The role of clinical care research evidence in clinical decisions [editorial]. *ACP Journal Club*, **125**, A14-A16.

HAYWARD, S. (1997). Strategies for regional development. *Local Economy*, **11**(4), 376-379.

HEAD, B. (2008). *Evidence-based policy: principles and requirements*. [online]. Last accessed 1 October 2014 at:
http://www.pc.gov.au/data/assets/pdf_file/0020/96203/roundtable-proceedings-volume1.pdf

HEALTH AND SOCIAL CARE INFORMATION CENTRE [HSCIC]. *Healthcare Resource Groups (HRG4)*. [online]. Last accessed 1 October 2014 at:
<http://www.hscic.gov.uk/hrg4>

HEALTH SERVICE JOURNAL [HSJ]. (2013). *Patient safety: the role and impact of clinical decision support*. [online]. Last accessed 23 October 2014 at:
<http://www.hsj.co.uk/patient-safety-the-role-and-impact-of-clinical-decision-support/5058746.article#.VFy3D8IKSfA>

HENTZELT, G. and LAUTERBACH, K. (2007). Can evidence-based medicine implicitly rely on current concepts of disease or does it have to develop its own definition? *Journal of Medical Ethics*, **33**(7), 394-399.

HIGGINS, J. and GREEN, S. (eds.). (2011). *Cochrane Handbook for Systematic Reviews of Interventions: Version 5.1*. [online]. Last accessed 23 October 2014 at: www.cochrane-handbook.org

HIGGINS, J., THOMPSON, S. and SPIEGELHALTER, D. (2009). A re-evaluation of random-effects meta-analysis. *Journal of the royal statistical society. Series A*, **172** (1), 137-159.

HIRASUNA, D. and HANSON, S. (2009). Introduction to the Special Issue. *International Regional Science Review*, **32**, 427-28.

HM GOVERNMENT (2012). *Civil Service Reform Plan*. London, HM Government.

- HM GOVERNMENT (2014). *Review of the Balance of Competences between the United Kingdom and the European Union Competition and Consumer Policy Report*. [online]. Last accessed 29 November 2007 at:
[https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/332779/Review of the Balance of Competences between the United Kingdom and the European Union.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/332779/Review_of_the_Balance_of_Competences_between_the_United_Kingdom_and_the_European_Union.pdf)
- HM TREASURY (2001). *Productivity in the UK: 3 – The Regional Dimension*. London, HMT.
- HM TREASURY (2003a). *A modern regional policy for the United Kingdom*. London, HMT, DTI ODPM.
- HM TREASURY (2003b). *Appraisal and Evaluation in Central Government*. London, HMSO.
- HM TREASURY (2004). *Devolving decision making 2: Meeting the regional economic challenge: increasing regional and local flexibility*. London, HMT.
- HM TREASURY (2007). *PSA Delivery Agreement 7: Improve the economic performance of all English regions and reduce the gap in economic growth rates between regions*. [online]. Last accessed 11 May 2014 at:
http://www.hmtreasury.gov.uk/d/pbr_csr07_psa7.pdf
- HM TREASURY (2011). *The Magenta Book: Guidance for Evaluation*. London, HMT.
- HM TREASURY PUBLIC EXPENDITURE STATISTICAL ANALYSES [HMT PESA] (2010). *Public Expenditure Statistical Analyses 2010*. [online]. Last accessed 29 July 2015 at:
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/220568/pesa_2010_complete.pdf
- HO, P., PETERSON, P. and MASOUDI, F. (2008). Evaluating the evidence: is there a rigid hierarchy? *Circulation*, **118**(16), 1675-1684.
- HOFFMAN, S. RØTTINGEN, A., BENNETT, S., LAVIS, J., EDGE, J. and FREN, J. (2012). *A Review of Conceptual Barriers and Opportunities Facing Health Systems Research to Inform a Strategy from the World Health Organization*. Geneva, World Health Organization.
- HOLFORD, N., KIMKO, H., MONTELEONE, J. and PECK, C. (2000). Simulation of clinical trials. *Annual Review of Pharmacology and Toxicology*, **40**, 209-234.
- HOUNTON, S. and NEWLANDS, D. (2012). Applying the net-benefit framework for assessing cost effectiveness of interventions towards universal health coverage. *Cost Effectiveness and Resource Allocation*, **10**, 8.
- HUBER, W. (2006). Evaluation of European Union Cohesion Policy: Window-dressing, Formal Exercise or Coordinated Learning Process? *Regional Studies*, **40**(2), 277-281.

HUGO, V. (1862). *Les Misérables*. Translated by Norman Denny. Reprint, New York, Penguin Books Ltd., 1982.

IPSOS-MORI (2012). *Review of 'Test, Learn, Adapt'* [online]. Last accessed 9 June 2014 at: https://www.ipsos-mori.com/Assets/Docs/Scotland/Approach%20Summer%202012/SRI_Scotland_Newsletter_Summer2011_Test_Learn_Adapt.pdf

ISSERMAN, A. and REPHANN, T. (1995) The economic effects of the Appalachian Regional Commission: An empirical assessment of 26 years of regional development planning. *Journal of the American Planning Association* **61**(3) 345-364.

JAFFE, A., TRAJTENBERG, M. and HENDERSON, R. (1993). Geographic localization of knowledge spillovers as evidenced by patent citations. *Quarterly Journal of Economics*, **79**, 577-598.

JISC (2009). *Demonstrating JISC's impact on the sector: Paper for HEFCE*. Annex A. [online]. Last accessed 1 October 2014 at: http://www.jisc.ac.uk/media/documents/committees/jir/2/jir_09_16iiscimpactreports_ep09annexa.pdf

JONES, H. (2009). *Policy making as discourse: a review of recent knowledge to policy literature*, Joint ODI- IKM Emergent Working paper. Bonn, EADI.

JONES, L. (1994). *The Social context of health and health work*. London, Palgrave Macmillan.

KERNER, J. F. (2006). Knowledge translation versus knowledge integration: A "funder's" perspective. *Journal of Continuing Education in the Health Professions*, **26** (1), 72–80.

KIEFER, L., FRANK, J., DI RUGGERIO, E., DOBBINS, M., MANUEL, D., GULLY, P. and MOWAT, D. (2005) Fostering Evidence-Based Decision-Making in Canada: Examining the Need for a Canadian Population and Public Health Evidence Centre and Research Network. *Canadian Journal of Public Health*, **96** (3) 11–119.

KITSON, M. (2004). Failure followed by success or success followed by failure? A Re-examination of British economic performance since 1950. In: FLOUD, R. and JOHNSON, P. (eds.). *The Cambridge Economic History of Britain*. Cambridge, Cambridge University Press.

KITSON, M. (2012). Regional competitiveness and regional policy: some insights from the UK. In: BRUGNOLI, A. and COLOMBO, A. (eds.) *Government, Governance and Welfare Reform: Structural Changes and Subsidiarity in Italy and Britain*. Cheltenham, Edward Elgar.

KNAAPEN, L. (2013). Being evidence-based in the absence of evidence: The management of non-evidence in guideline development. *Social Studies of Science*, **43**(5), 681-706.

- KRANE, V., ANDERSEN, M. and STREAN, W. (1997). Issues of qualitative research methods and presentations. *Journal of Sport and Exercise Psychology*, **19**, 213-218.
- KRUGMAN, P. (1991). Increasing Returns and Economic Geography. *The Journal of Political Economy*, **99** (3), 483-499.
- KUSHNER, S. (1996). The limits of constructivism. *Evaluation*, **2** (2), 189-200.
- LABOUR PARTY (1997). *New Labour: Because Britain Deserves Better*. London, The Labour Party.
- LAU, J., SCHMID, C. and CHALMERS, T. (1995). Cumulative meta-analysis of clinical trials builds evidence for exemplary medical care. *Journal of Clinical Epidemiology*, **48**, 45-57.
- LAVIS, J., ROBERTSON, D., WOODSIDE, J., MCLEOD, C. and ABELSON, J. (2003). How can research organizations more effectively transfer research knowledge to decision makers? *Milbank quarterly*, **81**, 221-248.
- LIN, V. (2003). Competing rationalities: evidence-based health policy. In: LIN, V. and GIBSON, B. (eds.) *Evidence-based health policy: problems & possibilities*. Oxford, Oxford University Press.
- LINCOLN, Y. and GUBA, E. (1985). *Naturalistic Inquiry*. Newbury Park CA, Sage Publications.
- LLOYD, P. and MEEGAN, R. (1996). Contested Governance: European Exposure in the English Regions. *European Planning Studies*, **4**, 75-97.
- LOCAL GOVERNMENT YORKSHIRE AND HUMBER SUB NATIONAL REVIEW OFFICERS GROUP. (2009). *Annex B: Terms of Reference and Membership of Regional Governance Structures*, 30 April 2009 [ONLINE]. Last accessed 3 July 2006 at: <http://www.publications.parliament.uk/pa/cm200809/cmselect/cm/york/memo/work/ucyh1202.htm#ftn5>
- LOCKE, E. (ed.) (1986). *Generalizing from laboratory to field settings*. Lexington MA, Lexington Books.
- LOMAS, J. (2007). The in-between world of knowledge brokering. *BMJ*, **334**(7585), 129-32.
- LSE GROWTH COMMISSION (2013). *Investing in Prosperity: Skills, Infrastructure and Innovation*. London, London School of Economics and Political Science.
- LUCAS, R. (1988). On the mechanics of economic development. *Journal of Monetary Economics*, **22**, 3-42.
- MARCH, J. (1978). Bounded rationality, ambiguity and the engineering of choice. *Bell Journal of Economics*, **9**, 587-608.

- MARCH, J. (1982). Theories of choice and making decisions. *Transaction-Society*, **20**, 29-39.
- MARCH, J. (1994). *A Primer on Decision Making: How Decisions Happen*. New York, Free Press.
- MARKUSEN, A. (2015). Problem-driven Research in Regional Science. *International Regional Science Review*, **38** (1), 3-29.
- MARSH, C. (1979). Problems with surveys: methods or epistemology. *Sociology*, **13** (2), 293-305.
- MARSTON, G. and WATTS, R. (2003). Tampering with the evidence: A critical appraisal of evidence-based policymaking. *Australian Review of Public Affairs*, **3**(3), 143-163.
- MARTIN, R., PIKE, A., TYLER, P. and GARDINER, B. (2015). *Spatially rebalancing the UK economy: the need for a new policy model*. [online]. Last accessed 29 July 2015 at: http://www.regionalstudies.org/uploads/documents/SRTUKE_v16_PRINT.pdf
- MARTIN, R. and MINNS, R. (1995). Undermining the Financial basis of regions: the spatial structure and implications of the UK pension system. *Regional Studies*, **29**, 12-144.
- MAY, T. (1997). *Social Research: issues, methods and process*. 2nd ed., Buckingham, Open University Press.
- MCCABE, C., CULYER, A. and CLAXTON, K. (2008). The NICE cost-effectiveness threshold: what it is and what it means. *Pharmacoeconomics*, **9**, 733-744.
- MCCAUGHEY, D. and BRUNING, N. (2010). Rationality versus reality: The challenges of evidence-based decision making for health policy makers. *Implementation Science*, **5**, 39-52.
- MCNAMARA, G., O'HARA, J., BOYLE, R. and SULLIVAN, C. (2009). Developing a culture of evaluation in the Irish public sector. *Evaluation: International Journal of Theory, Research & Practice*, **15**(1), 100-112.
- MCVITTIE, E. (2005). *The Information Requirements for an Effective Regional Policy*. Lecture presented to the Royal Society Statistics User Forum Conference: Neighbourhood and Regional Statistics. [lecture]. Held on 1 November, London UK.
- MCVITTIE, E. and SWALES, J. (2003). *Regional Policy Evaluation: Ignorance, Evidence and Influence*. Strathclyde Discussion Papers in Economics, No.03-06. [online]. Last accessed 3 July 2014 at: http://www.strath.ac.uk/media/departments/economics/researchdiscussionpapers/2003/media_34459_en.pdf

- MCVITTIE, E. and SWALES, J. (2007). The information requirements for an effective regional policy: a critique of the Allsopp report. *Urban Studies*, **44** (3), 425-438.
- MEDICAL RESEARCH COUNCIL (1948). Streptomycin treatment of pulmonary tuberculosis. *BMJ*, **2**, 769-782.
- MELTZER, D. (2001). Addressing uncertainty in medical cost-effectiveness analysis: Implications of expected utility maximization for methods to perform sensitivity analysis and the use of cost-effectiveness analysis to set priorities for medical research. *Journal of Health Economics*, **20** (1), 109-129.
- MENDEZ, C. (2011). The Lisbonization of EU Cohesion Policy: A Successful Case of Experimentalist Governance? *European Planning Studies*, **19** (3), 519-537.
- MISHAN, E., and QUAH, E. (2007). *Cost-Benefit Analysis*. UK, Routledge.
- MITTON, C., ADAIR, C. E., MCKENZIE, E., PATTEN, S. B. and WAYE PERRY, B. (2007). Knowledge transfer and exchange: review and synthesis of the literature. *The Milbank quarterly*, **85**, 729-68.
- MOLE, K., HART, M., ROPER, S. and SAAL, D. (2008). Differential gains from Business Link support and advice: A treatment effects approach. *Environment and Planning C, Government and Policy*, **26** (2), 315-334.
- MOORE, B. and RHODES, J. (1973). Evaluating the effects of British regional policy. *Economic Journal*, **83**, 87-110.
- MORGAN, K. (1997). The learning region: institutions, innovation and regional renewal. *Regional Studies*, **31**, 491-503.
- MURPHY, K. and FAFARD, P. (2012). Taking power, politics, and policy problems seriously: the limits of knowledge translation for urban health research. *Journal of Urban Health*, **89** (4), 723-32.
- NATIONAL AUDIT OFFICE [NAO] (2007). *Independent Performance Assessment: Yorkshire Forward*. [online]. Last accessed 7 April 2015 at: <http://www.nao.org.uk/wp-content/uploads/2007/03/Yorkshire-Forward2007.pdf>
- NATIONAL AUDIT OFFICE [NAO] (2010). *Independent Supplementary Reviews of the Regional Development Agencies: Summary Report*. [online]. Last accessed 1 October 2014 at: http://www.nao.org.uk/wp-content/uploads/2010/08/MDA_Independent_Reviews_rda.pdf
- NATIONAL AUDIT OFFICE [NAO] (2013a). *Evaluation in government*. [online]. Last accessed 1 October 2014 at: http://www.nao.org.uk/wp-content/uploads/2013/12/10331-001-Evaluation-in-government_NEW.pdf
- NATIONAL AUDIT OFFICE [NAO] (2013b). *Funding and structures for local economic growth*. HC 542 SESSION 2013-14. London, TSO.

- NATIONAL HEALTH SERVICE [NHS] (2014). *NHS Five Year Forward View*. [online]. Last accessed 29 July 2014 at: <http://www.england.nhs.uk/wp-content/uploads/2014/10/5yfv-web.pdf>
- NATIONAL HEALTH SERVICE ENGLAND [NHS ENGLAND] (2015). *Building the NHS of the Five year Forward View – NHS England Business Plan 2015/16*. [online]. Last accessed 29 July 2014 at: <http://www.england.nhs.uk/wp-content/uploads/2015/03/business-plan-mar15.pdf>
- NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE [NICE] (2013a). *NICE guidance: our programmes*. [online]. Last accessed 4 December 2014 at: <https://www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance>
- NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE [NICE] (2013b). *Guide to the Methods of Technology Appraisal* [online]. Last accessed 4 December 2014 at: <http://www.nice.org.uk/article/pmg9/chapter/foreword>
- NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE [NICE] (2013c). *Citizens Council Meeting Report: development of social care guidance*. [online]. Last accessed 4 December 2014 at: <https://www.nice.org.uk/news/press-and-media/nice-citizens-council-publishes-final-report-on-development-of-social-care-guidance>
- NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE [NICE] (2013d). *Social value judgements*. [online]. Last accessed 4 December 2014 at: <http://www.nice.org.uk/proxy/?sourceurl=http://www.nice.org.uk/aboutnice/howwe-work/socialvaluejudgements/socialvaluejudgements.jsp>
- NEWMAN, I. and BENZ, C. (1998). *Qualitative-quantitative research methodology: Exploring the interactive continuum*. Carbondale IL, Southern Illinois University Press.
- NHS CONNECTING FOR HEALTH EVALUATION PROGRAMME [NHS CFHEP]. (2011). *The Impact of eHealth on the Quality and Safety of Healthcare: An updated systematic overview & synthesis of the literature*. [online]. Last accessed 23 October 2014 at: <http://www.birmingham.ac.uk/Documents/college-mds/haps/projects/cfhep/projects/001Extension/CFHEP001FinalReport-March2011.pdf>
- NILSSON, M., JORDAN, A., TURNPENNY, J., HERTIN, J., NYKVIST, B. and RUSSEL, D. (2008). The Use and Non-Use of Policy Appraisal Tools in Public Policy Making: An Analysis of Three European Countries and the European Union. *Policy Sciences*, **41**, 335-355.
- NUTLEY, S. (2003). *Bridging the policy/ research divide: Reflections and Lessons from the UK*. Keynote paper for National Institute of Governance Conference. [lecture]. Held on 23/24 April, Canberra Australia.

NUTLEY, S. and WEBB, J. (2000). Evidence and the policy process. In: DAVIES, H., NUTLEY, S., and SMITH, P. (eds.) *What works? Evidence-based policy and practice in public services*. Bristol, The Policy Press, 13-41.

NUTLEY, S., DAVIES, H. and WALTER, I. (2002). *Evidence-based policy and practice: cross sector lessons from the UK*, Working Paper 9. St Andrews, University of St Andrews Research Unit for Research Utilisation.

NUTLEY, S., POWELL, A. and DAVIES H. (2013). *What Counts as Good Evidence?* London, Alliance for Useful Evidence.

NUTLEY, S., WALTER I. and DAVIES H. (2007). *Using evidence: How research can inform public services*. Bristol, Policy Press.

OFFICE FOR NATIONAL STATISTICS [ONS] (2011). *Measuring the economic impact of an intervention or investment, Paper Two: Existing sources and methods*. [online]. Last accessed 23 October 2014 at: <http://www.ons.gov.uk/ons/rel/regional-analysis/measuring-the-economic-impact-of-an-intervention-or-investment/index.html>

OFFICE OF PROJECT AND PROGRAMME ADVICE AND TRAINING [OffPAT]. *e-library*. [online]. Last accessed 1 January 2012 at: <http://www.nationalarchives.gov.uk/webarchive/central-government.htm#o>

OFFICE OF THE DEPUTY PRIME MINISTER [ODPM] (2004). *Assessing the Impacts of Spatial Interventions: Regeneration, Renewal and Regional Development*. 'The 3Rs guidance'. London, HMSO.

ONE NORTH EAST [ONE] (2012). *Annual Report and Accounts 2011-2012*. [online]. Last accessed 5 December 2014 at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/247067/0159.pdf

ONLINE LEARNING TASK FORCE [OLTF] (2011). *Collaborate to compete: Seizing the opportunity of online learning for UK higher education: Final report*. [online]. Last accessed 1 October 2014 at: http://www.hefce.ac.uk/media/hefce1/pubs/hefce/2011/1101/11_01.pdf

O'REILLY, D. (2007). Comment on Rhodes et al. (2005): Some Further Thoughts on Assessing the Effects of Area-based Initiatives on Local Outcomes. *Urban Studies*, **44** (5-6), 1145-1153.

OTTOSON, J. (2009). Knowledge-for-Action Theories in Evaluation: Knowledge Utilization, Diffusion, Implementation, Transfer, and Translation. *New Directions for Evaluation*, **124**, 7-20.

OVERMAN, H. (2013). *Alexandra Jones in conversation with Henry Overman: The effectiveness of current policy in stimulating economic growth in cities LSE Podcast*. [online] 13 March 2013. Last accessed 7 October 2014 at:

OVERMAN, H. and GIBBONS, S. (2011). Unequal Britain: How real are regional disparities? *Centrepiece*, **17** (1), 23-25.

PAPADAKIS, V. and BARWISE, P. (1997). What Can We Tell Managers About Making Strategic Decisions? In: PAPADAKIS, V. and BARWISE, P. (eds.). *Strategic Decisions*. London, Kluwer, 267-287.

PAPADAKIS, V., LIOUKAS, S. and CHAMBERS, D. (1998). Strategic Decision-Making Processes: The Role of Management and Context. *Strategic Management Journal*, **19** (2), 115-147.

PARSONS, W. (2002). From muddling through to muddling up-evidence based policy making and the modernisation of British Government. *Public Policy and Administration*, **17**, 43-60

PASCALL, G. (2007). Health and health policy. In: BALDOCK, J., MANNING, N. and VICKERSTAFF, S., (eds.) *Social policy*. 3rd ed., Oxford, Oxford University Press, 261-284.

PATTON, M. (1990). *Qualitative evaluation and research methods*. 2nd ed., Newbury Park CA, Sage.

PATTON, M. (2012). *Essentials of utilization-focused evaluation*. Thousand Oaks CA, Sage.

PAWSON, R. (2002). Evidence-based Policy: In Search of a Method. *Evaluation*, **8**, 2157-2181.

PAWSON, R. and TILLEY, N. (1997). *Realistic evaluation*. London, Sage.

PAWSON, R., GREENHALGH, T., HARVEY, G. and WALSHE, K. (2004). *Realist synthesis: an introduction*, ESRC Research Methods Programme Working Paper Series. [online]. Last accessed 1 October 2014 at:
<http://betterevaluation.org/sites/default/files/RMPmethods2.pdf>

PAWSON, R., GREENHALGH, T., HARVEY, G. and WALSHE, K. (2005). Realist review - a new method of systematic review designed for complex policy interventions. *Journal of Health Services Research & Policy*, **10**, 21-34.

PAWSON, R., BOAZ, A., GRAYSON, L., LONG, A. and BARNES, C. (2003). *Types and quality of knowledge in social care*, ESRC UK Centre for Evidence Based Policy and Practice. London, SCIE.

PERCY-SMITH, J., BURDEN, T., DARLOW, A., DOWSON, L., HAWTIN, M. and LADI, S. (2002). *Promoting change through research: the impact of research on local government*, Joseph Rowntree Foundation. [online]. Last accessed 1 October 2014 at:
<http://www.jrf.org.uk/sites/files/jrf/184263089x.pdf>

- PERVAN, G. and ARNOTT, D. (2005). A critical analysis of decision support systems research. *Journal of Information Technology*, **20**, 67-87.
- PETTICREW, M. and ROBERTS, H. (2003). Evidence, hierarchies, and typologies: horses for courses. *Journal of Epidemiology and Community Health*, **57**, 527-529.
- PETTICREW, M. and ROBERTS, H. (2006). *Systematic Reviews in the Social Sciences: A practical guide*. Oxford, Blackwell Publishing.
- PHILLIPS, C. (2009). *What is a QALY?* [online]. Last accessed 1 October 2014 at: <http://www.medicine.ox.ac.uk/bandolier/painres/download/whatis/qaly.pdf>
- PIKE, A., TOMANEY, J., COOMBES, M. and MCCARTHY, A. (2012). Governing uneven development: the politics of local and regional development in England. In: BELLINI, N., DANSON, M. and HALKIER, H. (eds.). *Regional Development Agencies: The Next Generation?* London, Routledge, 102-121.
- PIROG, M. (2014). Data will drive innovation in public policy and management research in the next decade. *Journal of Policy Analysis and Management*, **33**(2), 537-543.
- POLVERARI, L. and BACHTLER, J. (2004). *Assessing the Evidence: The Evaluation of Regional Policy in Europe*, Paper No. 04/5 to the European Regional Policy Research Consortium (EoRPA). Glasgow, University of Strathclyde European Policies Research Centre.
- PORTER, S. and CARTER, D. (2000). Common terms and concepts in research. In: CORMACK, D. (ed.). *The Research Process in Nursing*. Oxford, Blackwell Science.
- POWELL, M. (1997). *Evaluating the National Health Service*. Buckingham, Open University Press.
- PRESSMAN, J. and WILDAVSKY, A. (1984). *Implementation*. 3rd ed., Berkeley, University of California Press.
- PRICEWATERHOUSECOOPERS [PWC] (2009a). *Impact of RDA Spending National Report: Volume 1, Main Report*. London, PricewaterhouseCoopers.
- PRICEWATERHOUSECOOPERS [PWC] (2009b). *Impact of RDA Spending National Report: Volume 2, Regional Annexes*. London, PricewaterhouseCoopers.
- PUGA, D. (2010). The magnitude and causes of agglomeration economies. *Journal of Regional Science*, **50** (1), 203-219.
- PUGALIS, L. (2011). Sub-national economic development: Where do we go from here? *Journal of Urban Regeneration and Renewal*, **4** (3), 255-268.
- QUICK, C., REED, J., HARPER, S., SAEB-PARSY, K. and DEAKIN, P. (2013) *Essential surgery: problems, diagnosis and management*. 3rd ed, Edinburgh, Churchill Livingstone.

- RAHM, E. and DO, H. (2000). Data cleaning: Problems and current approaches. *IEEE Data Eng. Bull.* **23**(4), 3-13.
- RANDOLPH, A., HAYNES, R., WYATT, J., COOK, D. and GUYATT, G. (1999). Users' guides to the medical literature: XVIII. How to use an article evaluating the clinical impact of a computer-based clinical decision support system. *JAMA*, **282**, 67-74.
- REBELO, S. (1991). Long-run policy analysis and long-run growth. *The Journal of Political Economy*, **99** (3), 500-521.
- REED, M. and HARVEY, D (1992). The new science and the old: Complexity and realism in the social sciences. *Journal for the Theory of Social Behaviour*, **22**, 356-379.
- REEVES, B., DEEKS, J., HIGGINS, J. and WELLS, G. (2011). Including non-randomized studies. In: HIGGINS, J., GREEN, S. (eds.). *Cochrane Handbook for Systematic Reviews of Interventions: Version 5.0.1*. [online]. Last accessed 23 October 2014 at: www.cochrane-handbook.org
- RESEARCH COUNCILS UK [RCUK] (2014). *Local Economic Growth What Works Centre*. [online]. Last accessed 9 December 2014 at: <http://gtr.rcuk.ac.uk/project/431830CE-03C5-4BA0-AAE4-804C6A7B9133>
- RHODES, J., TYLER, P. and BRENNAN, A. (2005). Assessing the effects of area-based initiatives on local area outcomes: Some thoughts based on the National Evaluation of the Single Regeneration Budget in England. *Urban Studies*, **42** (11), 1919-1946.
- RITCHIE, J., LEWIS, J. and ELAM, G. (2003). Designing and selecting samples. In: RITCHIE, J. and LEWIS, J. (eds.). *Qualitative research practice: A guide for social science students and researchers*. London, Sage, 77-108.
- RITCHIE, J., SPENCER, L. and O'CONNOR, W. (2003). Carrying out qualitative analysis. In: RITCHIE, J. and LEWIS, J. (eds.). *Qualitative research practice: A guide for social science students and researchers*. London, Sage, 219-262.
- ROMER, P. M. (1986). Increasing returns and long-run growth. *Journal of Political Economy*, **94**, 1002-1037.
- RUTTER, A., HAWKINS, B. and PARKHURST, J. (2013). *Knowledge transfer and exchange: A look at the literature in relation to research and policy*, GRIP-Health Working Papers. London, London School of Hygiene and Tropical Medicine.
- RUTTER, J. (2012). *Evidence and evaluation in policy making: A problem of supply or demand?* Institute for Government. [online]. Last accessed 1 October 2014 at: <http://www.instituteforgovernment.org.uk/sites/default/files/publications/evidence%20and%20evaluation%20in%20template%20final%200.pdf>

SACKETT, D., ROSENBERG, W., GRAY, J., HAYNES, R, and RICHARDSON, W. (1996). Evidence-Based Medicine: What it is and what it isn't. *British Medical Journal*, **312**, 71-72.

SALTER, K. and KOTHARI, A. (2014). Using realist evaluation to open the black box of knowledge translation: a state-of-the-art review. *Implementation Science*, **9**, 115-129.

SANDERSON, I. (2002). Evaluation, policy learning and Evidence Based Policy. *Public Administration*, **80** (1), 1-22.

SANDERSON, I. (2003). Is it 'what works' that matters? Evaluation and evidence-based policy making. *Research Papers in Education*, **18** (4), 331-345.

SANDERSON, I. (2004). Getting evidence into practice: Perspectives on rationality. *Evaluation*, **10** (3), 366-379.

SANDERSON, I. (2006). Complexity, 'practical rationality' and evidence-based policy making. *Policy & Politics*, **34** (1), 115-132.

SAPSFORD, R. (1999). *Survey Research*. London, Sage.

SATTERFIELD, J., SPRING, B., BROWNSON, R., MULLEN, E., NEWHOUSE, R., WALKER, B. and WHITLOCK, E. (2009). Toward a Transdisciplinary Model of Evidence-Based Practice. *The Milbank Quarterly*, **87** (2), 368-390.

SAVAGE, L. (1954). *The Foundations of Statistics*. New York, Wiley.

SCHÖN, D.A. (1983). *The reflective practitioner: How professionals think in action*. New York, Basic books.

SCRIVEN, M. (1994). The fine line between evaluation and explanation. *Evaluation Practice*, **15**, 75-77.

SEFTON, T. (2000). *Getting less for more: Economic evaluation in the social welfare field*, CASEpaper 44. London, London School of Economics Political Science Centre for Analysis of Social Exclusion.

SEFTON, T. (2003). Economic evaluation in the social welfare field. *Evaluation*, **9**, 73-91.

SEFTON, T., BYFORD, S., MCDAID, D., HILLS, J. and KNAPP, M. (2002). *Making the most of it: Economic evaluation in the social welfare field*. York, York Publishing Services.

SHAXSON, L. (2005). Is your evidence robust enough? Questions for policy makers and practitioners. *Evidence and Policy: A Journal of Research, Debate and Practice*, **1**(1), 101-111.

SHAXSON, L., BIELAK, A., AHMED, I., BRIEN, D., CONANT, B., FISHER, C., GWYN, E., KLERKX, L., MIDDLETON, A., MORTON, S., PANT, L. and PHIPPS, D. (2012). *Expanding*

our understanding of K* (KT, KE, KTT, KMb, KB, KM, etc.). [online]. Last accessed 1 December 2014 at: http://inweh.unu.edu/wp-content/uploads/2013/05/KStar_ConceptPaper_FINAL_Oct29_WEB.pdf

SHERMAN, L., GOTTFREDSON, D., MACKENZIE, D., ECK, J., REUTER, P. and BUSHWAY, S. (1998). *Preventing crime: what works, what doesn't, what's promising*. Washington, USA National Institute of Justice.

SHOJANIA, K., SAMPSON, M., ANSARI, M., JI, J., DOUCETTE, S. and MOHER, D. (2007). How quickly do systematic reviews go out of date? A survival analysis. *Ann Intern Med.* **147**, 224-33.

SIMON, H. (1957). *Models of Man: Social and rational mathematical essays on rational human behaviour in a social setting*. New York, Wiley.

SKOCPOL, T. and SOMERS, M. (1980). The Uses of Comparative History in Macrosocial Inquiry. *Comparative Studies in Society and History*, **22**, 174-197.

SOLOW, R. (1956). A contribution to the theory of Economic Growth. *Quarterly Journal of Economics*, **70**, 65-94.

SOMEKH, B., BURMAN, E., DELAMONT, S., MEYER, J., PAYNE, M. and THORPE, R. (2005). Research communities in the social sciences. In: SOMEKH, B. and LEWIN, C. (eds.). *Research methods in the social services*. London, Sage, 1-14.

SORENSEN, C., DRUMMOND, M. and KANAVOS, P. (2008). *Ensuring value for money in health care: The role of health technology assessment in the European Union*. Copenhagen, WHO Regional Office for Europe.

SPENCE, D. (2014). Evidence based medicine is broken? *BMJ*, **348**, g22.

SPIEGELHALTER, D. and BEST, N. (2003). Bayesian approaches to multiple sources of evidence and uncertainty in complex cost-effectiveness modelling. *Statistics in Medicine*, **22**, 3687-3709.

SPIEGELHALTER, D., THOMAS, A., BEST, N. and LUNN, D. (2003). *WinBUGS user manual, version 1.4*. Cambridge, Medical Research Council Biostatistics Unit.

STERN, E. (2003). *The Rationale & Challenges for Evaluation of Public Policies*. Keynote Speech to the Fifth European Conference on Evaluation of the Structural Funds. [lecture]. Held on 26-27 June, Budapest.

STRAUS, S., TETROE, J. and GRAHAM, I. (2013) *Knowledge translation in health care: Moving from evidence to practice*. [online]. Last accessed 4 December 2014 at: <http://onlinelibrary.wiley.com/book/10.1002/9781118413555>

SUNSTEIN, C. (2000). Cognition and cost-benefit analysis. *The Journal of Legal Studies*, **29** (S2), 1059-1103.

- SUTCLIFFE S. and COURT, J. (2005). *Evidence-Based Policymaking: What is it? How does it work? What relevance for developing countries?* London, Overseas Development Institute.
- SUTTON, R. (1999). *The Policy Process: an Overview*, Working Paper 118. London, Overseas Development Institute.
- SWALES, K. (1997). A cost-benefit approach to the evaluation of regional selective assistance. *Fiscal Studies*, **18** (1), 73-85.
- SWIFT, B. (1996). Preparing numerical data. In: SAPSFORD, R. and JUPP, V. (eds.). *Data Collection and Analysis*. London, Sage.
- STUFFLEBEAM, D. L. and SHINKFIELD, A. J. (2007) *Evaluation theory, models, and applications*. San Francisco, Jossey-Bass.
- TATES, K., ZWAANSWIJK, M., OTTEN, RO., VAN DUMLLEN, S., HOOGERBRUGGE, P., KAMPS, W. and BENSING, J. (2009). Online focus groups as a tool to collect data in hard-to-include populations: examples from paediatric oncology. *BMC Medical Research Methodology*, **9**, 15.
- TAVISTOCK INSTITUTE (2003). *The Evaluation of Socio-Economic Development: The Guide*. London, Tavistock Institute.
- TAYLOR-GOOBY, P. and STOKER, G. (2011). The Coalition programme: a new vision for Britain or politics as usual? *The Political Quarterly*, **82** (1), 4-15.
- THE COCHRANE COLLABORATION (2014). *Evidence based health care and systematic reviews*. [online]. Last accessed 14 December 2014 at: <http://www.cochrane.org/about-us/evidence-based-health-care>
- TORGERSON, D. and TORGERSON, C. (2008). *Designing Randomized Trials in Health, Education and the Social Sciences*. Basingstoke UK, Palgrave Macmillan.
- TOWNSEND, A. (2012). The functionality of LEPs - are they based on travel to work? In: WARD, M. and HARDY, S. (eds.) (2012). *Changing Gear – Is Localism the New Regionalism?* London, The Smith Institute, 35-44.
- TYLER, P. and BRENNAN, A. (2007). Comments on Rhodes et al. (2005) - Some further thoughts on assessing the effects of area based local outcomes: A Reply. *Urban Studies*, **44** (5-6), 1155-1159.
- TYLER, P., WARNOCK, C., PROVINS, A. and LENZ, B. (2013). Valuing the benefits of urban regeneration. *Urban Studies*, **50**(1), 169-190.
- UNDERHILL, C. and OLMSTED, M. (2003). An experimental comparison of computer-mediated and face-to-face focus groups. *Social Science Computer Review*, **21** (4), 506-512.

- VAN DER KNAAP, P. (2006). Responsive evaluation and performance management: overcoming the downsides of policy objectives and performance indicators. *Evaluation*, **12** (3), 278-293.
- VENABLES, A. (2008). New Economic Geography. In: DURLAUF, S. and BLUME, L. (eds.). *The New Palgrave Dictionary of Economics*. 2nd ed., UK, Palgrave Macmillan.
- WAGSTAFF, A. and CULYER, A. (2012). Four decades of health economics through a bibliometric lens. *Journal of Health Economics*, **31** (2), 406-439.
- WARD, M. and HARDY, S. (eds.) (2012). *Changing Gear – Is Localism the New Regionalism?* London, The Smith Institute.
- WARD, V., HOUSE, A. and HAMER, S. (2009) Knowledge Brokering: The missing link in the evidence to action chain? *Evidence & policy : a journal of research, debate and practice*, **5** (3), 267-279.
- WEBSTER, C. (1998). *The National Health Service: A Political History*. Oxford, Oxford University Press.
- WEINSTEIN, M. and STASON, W. (1977). Foundations of cost-effectiveness analysis for health and medical practices. *New England Journal of Medicine*, **296** (13), 716-721.
- WEISS, C. (1979). The many meanings of research utilization. *Public Administration Review*, **39**, 426-431.
- WEISS, M., BLOOM, H. and BROCK, T. (2014). A Conceptual Framework for Studying the Sources of Variation in Program Effects. *Journal of Policy Analysis and Management*, **33** (3), 778-808.
- WELLS, P. (2007). New Labour and evidence based policy making: 1997-2007. *People, Place & Policy Online*, **1** (1), 22-29.
- WELLS, P. (2014). Professor of Public Policy Analysis and Evaluation, CRESR. Conversation with the author, 15th October. Personal communication.
- WHAT WORKS CENTRE FOR LOCAL ECONOMIC GROWTH [WWG] (2014). *Evidence Review: Employment Training*. [online]. Last accessed 3 July 2006 at: <http://whatworksgrowth.org/wp-content/uploads/2014/04/Employment-Training-Full-Report.pdf>
- WHITE, M. (1994). The gift of tired tongues: Gordon Brown's Speech at an economic seminar, Tuesday 27 September. *The Guardian*, 30 September.
- WILSON, I. (2012). Using Shadow Pricing to Value Outcomes from Regeneration Programmes: Evidence from the New Deal for Communities Programme in England. *The Town Planning Review*, **83** (6), 669-694.

WILSON, I (2013). Outcomes for 'stayers' in urban regeneration areas: the New Deal for Communities Programme in England. *Urban Research & Practice*, 6(2), 174-193.

WORRALL, J. (2010). Evidence: philosophy of science meets medicine. *Journal of Evaluation in Clinical Practice*, 16, 356-362.

WREN C. and TAYLOR J. (1999). Industrial restructuring and regional policy. *Oxford Economic Papers*, 51, 487-516.

YORKSHIRE FORWARD [YF] (2009). *Yorkshire Forward* [online]. Last accessed 5 January 2015 at:
<http://webarchive.nationalarchives.gov.uk/20091205142332/http://www.yorkshire-forward.com/>

YORKSHIRE FORWARD [YF] (2010). *YF Policy Product Range Evaluations – Estimating Potential GVA by Regeneris Consulting*. Unpublished.

YORKSHIRE FORWARD [YF] (2011a). *Annual Report and Accounts 2010-2011*. [online]. Last accessed 5 December 2014 at:
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/247235/1386.pdf

YORKSHIRE FORWARD [YF] (2011c). *Yorkshire Forward and Economic Development. Learning Legacy Module 1: Economic Strategy* [online]. Last accessed 5 January 2015 at: <http://www.cdialliance.co.uk/wp-content/uploads/2011/08/1-Economic-Strategy-final.pdf>

YORKSHIRE FORWARD [YF] (2011c). *Yorkshire Forward and Economic Development. Learning Legacy Module 2: Research, Intelligence and Evaluation* [online]. Last accessed 5 January 2015 at: <http://www.cdialliance.co.uk/wp-content/uploads/2011/08/2-Research-Intelligence-and-Evaluation-final.pdf>

YORKSHIRE FORWARD [YF] (2012). *Annual Report and Accounts 2011-2012*. [online]. Last accessed 5 December 2014 at:
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/246993/0319.pdf

YOUNG, E. and QUINN L. (2002). *Writing effective public policy papers: A guide to policy advisers in Central and Eastern Europe*. Budapest, LGI.

ZWARENSTEIN, M. and TREWEEK, S. (2009). What kind of randomized trials do we need? *Journal of Clinical Epidemiology*, 62, 461-463.

WORD COUNT: 80,560 (89,047 including the Bibliography)

Sheffield Hallam University
Learning and Information Services
Adsetts Centre, City Campus
Sheffield S1 1WD

28102

102 144 999 7



Sheffield Hallam University
Learning and Information Services
Adsetts Centre, City Campus
Sheffield S1 1WD

REFERENCE

Appendix

Appendix 1: Example of RDA programmes, projects and links to IEF sub-themes

Table 1: Yorkshire Forward – an example of RDA programmes, projects and links to IEF sub-themes

Policy Product Range (programme level)	Product (project level)	IEF sub-theme
Competitiveness	Business Improvement	Sector/cluster support
	Innovation	Science, R&D and innovation infrastructure
Economic Infrastructure	Delivering Sustainable Construction	Sustainable consumption/production
	Developing effectiveness of URCS & EDCs	Cross-cutting regeneration initiatives
	Funding and/or enabling infrastructure	Cross-cutting regeneration initiatives
	Maximise PSI, expertise & EU funding	Cross-cutting regeneration initiatives
	Influence the Region's major utilities	Other Regeneration initiatives
	Restoration & Redev of former Coalfield Sites	Bringing land back into use
Employment	Equality and Diversity	Matching people to jobs
	Increasing the No of People into Employment	Matching people to jobs
	Local economic dev in the more deprived areas	Individual Enterprise Level Support
Enterprise/Access to Finance	Access to Finance	Individual Enterprise Level Support
	Business Start Up Programme	Individual Enterprise Level Support
International Business	Stimulating an Enterprise Culture	Skills and workforce development
	Foreign Direct Investment	Inward investment
	Investor Development	Inward investment
	Trade Activity	Internationalisation of indigenous business
Low Carbon Economy	Energy and Climate Change	Sector/cluster support

	Environmental Assets	Other Regeneration initiatives
	Low Carbon Technologies	Sector/cluster support
	Sustainable Consumption and Production	Sustainable consumption/production
Major Events / Culture / Tourism	Marketing the Region through Major Events	Image. Events and tourism
	Visitor Economy	Image. Events and tourism
Rural Renaissance	Rural Capitals	Cross-cutting regeneration initiatives
	Rural Enterprise	Individual Enterprise Level Support
	Rural Policy	Other Regeneration initiatives
	Rural services	Individual Enterprise Level Support
Skills	Developing Skills Solutions - Higher Level Skills	Skills and workforce development
	Raising Aspirations of Individuals	Skills and workforce development
Transport	Stimulate Business Demand	Skills and workforce development
	Achieving consensus on transport priorities	Cross-cutting regeneration initiatives
	Developing Evidence	Cross-cutting regeneration initiatives
	Influence Transport Providers and Operators	Cross-cutting regeneration initiatives
	Work with partners to invest in transport	Cross-cutting regeneration initiatives
Urban Renaissance	Learning and Growing Places	Cross-cutting regeneration initiatives
	Making Places	Cross-cutting regeneration initiatives
	Thinking places	Cross-cutting regeneration initiatives
	Designing Places	Cross-cutting regeneration initiatives

Source: YF (2010).

Appendix 2: Survey items and respondents - online survey

Table 1: Survey items and respondents overview – online survey

Area	Respondents	No.	Question	Format	Response options
Identifier	All	1	Name (kept anonymous)	Open-ended	
Characteristics	RDA	2	Regional Development Agency (RDA)	Open-ended	
		3	Job title (kept anonymous)	Open-ended	
		4	Time worked at the RDA	Matrix of drop-down menus	From: 1999 To: 2011
		5	Which department did you work for when you were involved in Regional policy evaluation?	Open-ended	
	BIS/CLG	6	Job title (kept anonymous)	Open-ended	
		7	Which consultancy firm did you work for when you produced [RDA evaluation] work? (kept anonymous)	Open-ended	
		8	Job title (kept anonymous)	Open-ended	
		9	What role have you played in the evaluation of Regional policy and the evaluation of the Regional Development Agencies (RDAs) specifically?	Open-ended	
Verifying knowledge of evaluation processes	BIS/CLG				
	RDA delivery/performance	10	Were you involved in any evaluations at	Open-	

Overall influence of evaluation			the RDA? If so, which evaluations were you involved with (please include project numbers) and what was your role?	ended	
	Evaluators		11 Which RDA have you produced most work for?	Open-ended	
	All		12 (a) Please indicate your level of agreement or disagreement with the following statements concerning the role of evaluation within the RDA.	Matrix of choices (single select per row)	<p><u>Row choices:</u></p> <ul style="list-style-type: none"> - Clarified objectives and strategies for decision making - Provided evidence for economic appraisal of individual projects - Provided evidence for RDA investment prioritisation - Improved project design and development - Increased management and delivery efficiency - Enhanced performance and effectiveness - Showcased the effectiveness of projects to internal and external audiences - Highlighted "what works" and "what does not

Evaluation team processes					work” - Increased the capacity to conduct critical self-assessment - Contributed to the evidence base - It was a BIS requirement - It was an EU requirement - Evaluation was not important - Too much money was spent on evaluation <u>Column choices:</u> - Strongly agree - Agree - Neither agree or disagree - Disagree - Strongly disagree - Don't know
	BIS/CLG & evaluators	12 (b)	Other (please specify) / further comments	Open-ended	
		12 (c)	Was there any variation across the RDAs?	Open-ended	
		13 (a)	How would you rate the evaluation team processes at the RDAs?	Matrix of choices (single select per	<u>Row choices:</u> - Planning and preparing for evaluation - Committing RDA
	All				

				row)	<p>resources and support for evaluation</p> <ul style="list-style-type: none"> - Managing evaluation cost-effectively - Ensuring evaluation was carried out robustly - Ensuring evaluation met the guidelines of the RDA - Ensuring evaluation met the guidelines of BIS - Ensuring evaluation met the guidelines of the EU - Ensuring transparency in the conduct of evaluation - Ensuring independence in the conduct of evaluation - Ensuring inclusiveness in the conduct of evaluation - Disseminating the results of evaluation within the RDA - Disseminating the results of evaluation to external audiences - Ensuring learning and development from evaluation findings
--	--	--	--	------	--

Challenges faced by evaluation							<u>Column choices:</u> - Very good - Good - Neither good nor poor - Poor - Very poor - Don't know
				13 (b)	Other (please specify) / further comments	Open-ended	
				14	Thinking about your answers above, how do you think evaluation team processes could have been improved and why?	Open-ended	
				15	What do you think were the main successes and failures of the evaluation team?	Open-ended	
				16	Was there any variation across the RDAs?	Open-ended	
Challenges faced by evaluation	All			17 (a)	What strategic, operational and relational challenges do you think the evaluation team faced?	Matrix of choices (single select per row)	<u>Row choices:</u> - Strategic: Role of evaluation and evaluation processes were not clearly defined within RDA strategic documents (RES etc) - Strategic: Changing guidelines and expectations from BIS

					<ul style="list-style-type: none"> - Strategic: Changing guidelines and expectations from senior management within the RDA - Operational: Lack of budget for evaluations - Operational: Lack of budget for evaluation team staff - Operational: Lack of budget for evaluation team training and development - Operational: Lack of budget for dissemination activities - Relational: Lack of senior management support for evaluation - Relational: Lack of project manager support for evaluation - Relational: Lack of external partner's support for evaluation
Column choices:					

					<ul style="list-style-type: none"> - Yes - No - Don't know
	17 (b)	Other (please specify) / further comments	Open-ended		
	18	Thinking about your responses to the question above, how do you think the consequences of these challenges played out?	Open-ended		
	19	Was there any variation across the RDAs?	Open-ended		
Evaluation influence on strategic decision making	BIS/CLG & evaluators				
	20 (a)	Please indicate your level of agreement or disagreement with the following statements concerning the role of evaluation to inform strategic decision making within the RDA.	Matrix of choices (single select per row)	<p><u>Row choices:</u></p> <ul style="list-style-type: none"> - Evaluation evidence fed into the Corporate Plan - Evaluation evidence fed into the RES - Evaluation evidence fed into the Integrated Regional Strategy - Evaluation evidence fed into the ERDF Programme strategy - Evaluation evidence fed into decision making at Board meetings - Evaluation evidence fed into decision making at Executive meetings 	

						<ul style="list-style-type: none"> - Evaluation evidence informed the investment prioritisation exercise due to the Government's budget cuts - Evaluation evidence did not inform strategic decision making
						<u>Column choices:</u> <ul style="list-style-type: none"> - Strongly agree - Agree - Neither agree or disagree - Disagree - Strongly disagree - Don't know
						Open-ended
						Open-ended
						Open-ended
Evaluation influence on investment prioritisation	BIS/CLG & evaluators		20 (b)	Other (please specify) / further comments	Open-ended	
			21	How do you think evaluation evidence could/should have informed strategic decision making at the RDA?	Open-ended	
			22	Was there any variation across the RDAs?	Open-ended	
			23	In general, how was investment prioritisation undertaken at the RDA?	Open-ended	
Evaluation influence on investment	RDA evaluation team & strategy team		24	Specifically, how was the Coalition's budget cuts (announced in 2010)	Open-ended	

prioritisation (case study = budget cuts)			implemented at the RDA?		
		25	How was evaluation evidence used to inform the budget cuts investment prioritisation exercise?	Open-ended	
	RDA evaluation team & strategy team & evaluators	26	What evaluation evidence would have been useful to inform the budget cuts investment prioritisation exercise?	Open-ended	
Budget cuts	Strategy team	27 (a)	How important were each of the following challenges for the RDA when implementing the budget cuts in 2010?	Matrix of choices (single select per row)	<p><u>Row choices:</u></p> <ul style="list-style-type: none"> - Lack of time - Lack of resources/staff - Lack of clarity over the level of budget reduction - Lack of clarity over strategic priorities from BIS - Lack of clarity over strategic priorities from RDA senior management - Lack of clarity over strategic priorities from partners - Lack of monitoring data - Robustness of monitoring data - Lack of evaluation evidence - Robustness of evaluation evidence

				row)	projects - Detailed review of alternative options - Expected value for money assessment - Assessment of proposals of management and delivery of the project <u>Column choices:</u> - Very good - Good - Neither good nor poor - Poor - Very poor
29 (b)		Other (please specify) / further comments	Open-ended		
30		How do you think economic appraisal could/should have been conducted at the RDA?	Open-ended		
31 (a)		Were benchmarks (regional or national) used to inform economic appraisal?	Multiple choice (single select)	- Yes - No	
31 (b)		If not, why not?	Open-ended		
32 (a)		Was Cost Benefit Analysis used for economic appraisal?	Multiple choice	- Yes - No	

					(single select)	
		32 (b)	If not, why not? (i.e. was CBA applied differently for different projects?)	Open-ended		
Evaluation influence on economic appraisal	RDA evaluation team & economic appraisal team & strategy team & BIS/CLG & evaluators	33 (a)	Did evaluation evidence inform economic appraisal at the RDA?	Multiple choice (single select)	- Yes - No	
		33 (b)	If yes, how?	Open-ended		
		34	How do you think evaluation evidence could/should have been used to inform economic appraisal at the RDA?	Open-ended		
Evaluation influence on project delivery	BIS/CLG & evaluators	35	Was there any variation across the RDAs?	Open-ended		
		36 (a)	Did evaluation inform project delivery at the RDA?	Multiple choice (single select)	- Yes - No	
		36 (b)	If yes, how?	Open-ended		
		37	How do you think evaluation could/should have informed project delivery at the RDA?	Open-ended		
Monitoring	BIS/CLG & evaluators	38	Was there any variation across the RDAs?	Open-ended		
	Delivery/performance team	39	How were projects monitored at the RDA?	Open-ended		

Evaluation strategy					on spend and outputs into a Management Information System	<ul style="list-style-type: none"> - Inputting beneficiary data into a Management Information System - Reporting and analysing monitoring data for internal audiences - Disseminating monitoring data for external audiences 	<p><u>Column choices:</u></p> <ul style="list-style-type: none"> - Very good - Good - Neither good nor poor - Poor - Very poor
Evaluation strategy	RDA evaluation team & BIS/CLG & evaluators		43 (b)	Other (please specify) / further comments	Open-ended		
			44	How were projects and programmes evaluated at the RDA(s)?	Open-ended		
			45 (a)	Was there a focus on project evaluation or programme level evaluation?	Multiple choice (single select)	<ul style="list-style-type: none"> - Project - Programme - Project and programme - Neither 	
			45	What were the reasons for this?	Open-		

Impact of central government evaluation processes		(b)		ended	
		46	How do you think projects and programmes could/should have been evaluated at the RDA(s)?	Open-ended	
	BIS/CLG & evaluators	47	Was there any variation across the RDAs?	Open-ended	
	RDA evaluation team & BIS/CLG & evaluators	48 (a)	Did the introduction of the Impact Evaluation Framework (IEF) have the following implications?	Matrix of choices (single select per row)	<u>Row choices:</u> <ul style="list-style-type: none"> - More evaluations conducted - More money spent on evaluation in total - More money spent per individual evaluation conducted - More money spent per individual evaluation conducted - More focus on quantitative data - More focus on primary surveys over secondary data use - More focus on Strategic Added Value - More focus on social impacts - More focus on geographical impacts

				<ul style="list-style-type: none"> - More focus on environmental impacts - More focus on qualitative data - More focus on standardised evaluation reports which were less illuminating - More focus on the synthesis of evaluations - More focus on the use of benchmarks (regional and national) - More focus on final evaluations - More focus on interim evaluations and forecast impact
				<p><u>Column choices:</u></p> <ul style="list-style-type: none"> - Strongly agree - Agree - Neither agree or disagree - Disagree - Strongly disagree
48 (b)	Other (please specify) / further comments	Open-ended		
49	Did the PWC exercise to collate	Matrix of		<u>Row choices:</u>

		(a)	evaluation evidence for the 'Impact of RDA spending: national report' for BERR have the following implications?	choices (single select per row)	<ul style="list-style-type: none"> - More evaluations conducted - More money spent on evaluation in total - More money spent per individual evaluation conducted - More focus on quantitative data - More focus on primary surveys over secondary data use - More focus on Strategic Added Value - More focus on social impacts - More focus on geographical impacts - More focus on environmental impacts - More focus on qualitative data - More focus on standardised evaluation reports which were less illuminating - More focus on the synthesis of evaluations
--	--	-----	---	---------------------------------	---

Value for money of evaluation					<ul style="list-style-type: none">- More focus on the use of benchmarks (regional and national)- More focus on final evaluations- More focus on interim evaluations and forecast impact	
					<u>Column choices:</u> <ul style="list-style-type: none">- Strongly agree- Agree- Neither agree or disagree- Disagree- Strongly disagree	
				49 (b)	Other (please specify) / further comments	Open-ended
	BIS/CLG & evaluators			50	Was there any variation across the RDAs?	Open-ended
	All			51 (a)	In your opinion do you think conducting evaluations is good value for money?	Multiple choice (single select)
			51 (b)	Can you please explain your response?	Open-ended	
			52	Thinking about your answer above, how do you think value for money can be improved?	Open-ended	
					<ul style="list-style-type: none">- Yes - in all cases- Yes - in some cases- No	

Evidence synthesis & future evaluation	RDA evaluation team & BIS/CLG & evaluators	53 (a)	What do you think are the main challenges for synthesising the current evaluation evidence base to inform decision making for the LEPs and the Regional Growth Fund?	Multiple choice likert scale	<p><u>Row choices:</u></p> <ul style="list-style-type: none"> - Difficulty in synthesising qualitative data - Difficulty in capturing the importance of context for each individual evaluation - Lack of robust monitoring data - IEF has not been applied in the same way across evaluations - Some evaluations are not IEF compliant - Some evaluations do not explicitly note the additionality calculation (deadweight, displacement, substitution, multipliers) - Differences in objectives of projects and programmes - Selection bias of the evaluation reports (more likely to show the RDA in a positive light) - Poorly articulated logic chains and theories of
--	--	--------	--	------------------------------	---

Evaluation methodology		53 (b)	Other (please specify) / further comments	Open-ended	change Column choices: - Strongly agree - Agree - Neither agree or disagree - Disagree - Strongly disagree
			What do you think the future is for Regional policy evaluation (LEPs/Regional Growth Fund)?	Open-ended	
			What factors do you think could impact on the future of Regional policy evaluation?	Open-ended	
	RDA evaluation team (in-depth questionnaire)	54	What primary data collection and analysis methods were used for evaluations at the RDA? How do you think the use of primary data could have been improved?	Open-ended	
		56	What secondary data collection and analysis methods were used for evaluations at the RDA? How do you think the use of secondary data could have been improved?	Open-ended	
		57	How was the counterfactual (what would have happened otherwise)	Open-ended	

		identified, calculated and quality checked for evaluations at the RDA? How do you think this could have been improved?		
59		How were robust sample sizes for primary data collection determined and quality checked at the RDA? How do you think this could have been improved?	Open-ended	
60		How was additionality (deadweight, displacement, substitution, multipliers) calculated and quality checked for evaluations at the RDA? How do you think this could have been improved?	Open-ended	
61		How was Value for Money and Return on Investment calculated and quality checked at the RDA? How do you think this could have been improved?	Open-ended	
62		It could be argued that RDA commissioned evaluations were more likely to be supportive of RDA interventions, leading to bias. How was this accounted for at the RDA? How do you think this could have been improved?	Open-ended	
63		How were evaluation findings tested to produce recommendations at the RDA? How do you think this could have been improved?	Open-ended	

Overall		64	How were evaluation findings disseminated internally? How do you think this could have been improved?	Open-ended	
		65	How were evaluation findings disseminated externally? How do you think this could have been improved?	Open-ended	
	All	66	Please use the space below for any additional comments.	Open-ended	

Appendix 3: Online survey respondents by RDA

Table 1: Online survey respondents by organisation

Respondent's organisation	Number
RDA:	
- YF	48
- ONE	6
- SWRDA	6
- SEEDA	1
- LDA	1
- NWDA	1
- EMDA	1
- AWM	1
External consultancy	13
Central government	3
TOTAL (N)	81

Appendix 4: Example of RDA staff numbers by department

Table 1: emda staff numbers by departmental sub-group

	2011-2012 staff nos.	Restated 2010-2011 staff nos.
Chief Executive Department	4	5
Corporate Services	22	52
Regeneration Development Directorate	8	55
Strategy and Communications Directorate	4	44
Business Services Directorate	10	72
	<u>48</u>	<u>228</u>

Source: emda Annual Report and Accounts 2011-12, p73

Table 1 gives an example of RDA staff numbers by department for the East Midlands Development Agency (emda). This demonstrates that the strategy and communications directorate (which included evaluation and strategy teams) accounted for less than 20 per cent of the RDA's staff in 2010-11 and less than 10 percent in 2011-2012.

Appendix 5: Quantitative survey results

Table 1: Overall influence of evaluation in the RDAs – total population¹¹⁶

Question 12(a): Please indicate your level of agreement or disagreement with the following statements concerning the role of evaluation within the RDA.	Strongly agree	Agree	Neither agree or disagree	Disagree	Strongly disagree	Don't know	Response Count
Clarified objectives and strategies for decision making	6	54	11	7	2	1	81
Provided evidence for economic appraisal of individual projects	19	54	1	6	0	1	81
Provided evidence for RDA investment prioritisation	10	42	18	9	1	1	81
Improved project design and development	9	40	21	6	2	3	81
Increased management and delivery efficiency	6	31	27	11	2	4	81
Enhanced performance and effectiveness	7	37	26	7	2	2	81
Showcased the effectiveness of projects to internal and external audiences	23	45	10	0	0	3	81
Highlighted “what works” and “what does not work”	24	45	8	2	0	2	81
Increased the capacity to conduct critical self-assessment	11	37	19	9	2	3	81
Contributed to the evidence base	28	48	2	2	0	1	81
It was a BIS requirement	31	30	11	3	0	6	81
It was an EU requirement	25	33	9	6	0	8	81
Evaluation was not important	3	6	7	34	29	2	81
Too much money was spent on evaluation	2	16	13	32	15	3	81

¹¹⁶ For reasons of clarity, in the narrative the combined results of ‘strongly agree and ‘agree’ are combined to express agreement.

Table 2: Overall influence of evaluation in the RDAs – total population (%)

Question 12(a): Please indicate your level of agreement or disagreement with the following statements concerning the role of evaluation within the RDA.	Strongly agree	Agree	Neither agree or disagree	Disagree	Strongly disagree	Don't know	Response Count
Clarified objectives and strategies for decision making	7%	67%	14%	9%	2%	1%	81
Provided evidence for economic appraisal of individual projects	23%	67%	1%	7%	0%	1%	81
Provided evidence for RDA investment prioritisation	12%	52%	22%	11%	1%	1%	81
Improved project design and development	11%	49%	26%	7%	2%	4%	81
Increased management and delivery efficiency	7%	38%	33%	14%	2%	5%	81
Enhanced performance and effectiveness	9%	46%	32%	9%	2%	2%	81
Showcased the effectiveness of projects to internal and external audiences	28%	56%	12%	0%	0%	4%	81
Highlighted “what works” and “what does not work”	30%	56%	10%	2%	0%	2%	81
Increased the capacity to conduct critical self-assessment	14%	46%	23%	11%	2%	4%	81
Contributed to the evidence base	35%	59%	2%	2%	0%	1%	81
It was a BIS requirement	38%	37%	14%	4%	0%	7%	81
It was an EU requirement	31%	41%	11%	7%	0%	10%	81
Evaluation was not important	4%	7%	9%	42%	36%	2%	81
Too much money was spent on evaluation	2%	20%	16%	40%	19%	4%	81

Table 3: Overall influence of evaluation in the RDAs – YF sub-population

Question 12(a): Please indicate your level of agreement or disagreement with the following statements concerning the role of evaluation within the RDA.	Strongly agree	Agree	Neither agree or disagree	Disagree	Strongly disagree	Don't know	Response Count
Clarified objectives and strategies for decision making	4	33	6	3	2	0	48
Provided evidence for economic appraisal of individual projects	11	33	1	3	0	0	48
Provided evidence for RDA investment prioritisation	6	24	12	5	1	0	48
Improved project design and development	8	21	14	3	1	1	48
Increased management and delivery efficiency	4	24	12	5	2	1	48
Enhanced performance and effectiveness	4	25	14	3	2	0	48
Showcased the effectiveness of projects to internal and external audiences	12	30	4	0	0	2	48
Highlighted “what works” and “what does not work”	15	26	5	1	0	1	48
Increased the capacity to conduct critical self-assessment	7	26	8	4	2	1	48
Contributed to the evidence base	14	30	2	2	0	0	48
It was a BIS requirement	13	21	6	3	0	5	48
It was an EU requirement	16	19	5	2	0	6	48
Evaluation was not important	2	3	3	16	23	1	48
Too much money was spent on evaluation	1	10	6	20	9	2	48

Table 4: Overall influence of evaluation in the RDAs – YF sub-population (%)

Question 12(a): Please indicate your level of agreement or disagreement with the following statements concerning the role of evaluation within the RDA.	Strongly agree	Agree	Neither agree or disagree	Disagree	Strongly disagree	Don't know	Response Count
Clarified objectives and strategies for decision making	8%	69%	13%	6%	4%	0%	48
Provided evidence for economic appraisal of individual projects	23%	69%	2%	6%	0%	0%	48
Provided evidence for RDA investment prioritisation	13%	50%	25%	10%	2%	0%	48
Improved project design and development	17%	44%	29%	6%	2%	2%	48
Increased management and delivery efficiency	8%	50%	25%	10%	4%	2%	48
Enhanced performance and effectiveness	8%	52%	29%	6%	4%	0%	48
Showcased the effectiveness of projects to internal and external audiences	25%	63%	8%	0%	0%	4%	48
Highlighted “what works” and “what does not work”	31%	54%	10%	2%	0%	2%	48
Increased the capacity to conduct critical self-assessment	15%	54%	17%	8%	4%	2%	48
Contributed to the evidence base	29%	63%	4%	4%	0%	0%	48
It was a BIS requirement	27%	44%	13%	6%	0%	10%	48
It was an EU requirement	33%	40%	10%	4%	0%	13%	48
Evaluation was not important	4%	6%	6%	33%	48%	2%	48
Too much money was spent on evaluation	2%	21%	13%	42%	19%	4%	48

Table 5: The effectiveness of RDA evaluative research processes – total population¹¹⁷

Question 13(a): How would you rate the evaluation team processes at the RDAs?	Very good	Good	Neither good nor poor	Poor	Very poor	Don't know	Response Count
Planning and preparing for evaluation	27	41	6	3	0	2	79
Committing RDA resources and support for evaluation	13	51	10	2	0	4	80
Managing evaluation cost-effectively	12	41	17	4	1	5	80
Ensuring evaluation was carried out robustly	28	35	10	5	0	2	80
Ensuring evaluation met the guidelines of the RDA	30	43	4	0	0	3	80
Ensuring evaluation met the guidelines of BIS	24	37	9	2	0	8	80
Ensuring evaluation met the guidelines of the EU	26	30	12	1	0	11	80
Ensuring transparency in the conduct of evaluation	17	46	10	4	0	3	80
Ensuring independence in the conduct of evaluation	26	41	6	2	1	4	80
Ensuring inclusiveness in the conduct of evaluation	9	46	16	2	0	7	80
Disseminating the results of evaluation within the RDA	4	27	20	20	2	7	80
Disseminating the results of evaluation to external audiences	2	15	29	18	5	11	80
Ensuring learning and development from evaluation findings	2	26	27	19	1	6	81

¹¹⁷ For reasons of clarity, in the narrative the combined results of 'very good' and 'good' are combined to express agreement.

Table 6: The effectiveness of RDA evaluative research processes – total population (%)

Question 13(a): How would you rate the evaluation team processes at the RDAs?		Very good	Good	Neither good nor poor	Poor	Very poor	Don't know	Response Count
Planning and preparing for evaluation		34%	52%	8%	4%	0%	3%	79
Committing RDA resources and support for evaluation		16%	64%	13%	3%	0%	5%	80
Managing evaluation cost-effectively		15%	51%	21%	5%	1%	6%	80
Ensuring evaluation was carried out robustly		35%	44%	13%	6%	0%	3%	80
Ensuring evaluation met the guidelines of the RDA		38%	54%	5%	0%	0%	4%	80
Ensuring evaluation met the guidelines of BIS		30%	46%	11%	3%	0%	10%	80
Ensuring evaluation met the guidelines of the EU		33%	38%	15%	1%	0%	14%	80
Ensuring transparency in the conduct of evaluation		21%	58%	13%	5%	0%	4%	80
Ensuring independence in the conduct of evaluation		33%	51%	8%	3%	1%	5%	80
Ensuring inclusiveness in the conduct of evaluation		11%	58%	20%	3%	0%	9%	80
Disseminating the results of evaluation within the RDA		5%	34%	25%	25%	3%	9%	80
Disseminating the results of evaluation to external audiences		3%	19%	36%	23%	6%	14%	80
Ensuring learning and development from evaluation findings		2%	32%	33%	23%	1%	7%	81

Table 7: The effectiveness of RDA evaluative research processes – YF sub-population

Question 13(a): How would you rate the evaluation team processes at the RDAs?	Very good	Good	Neither good nor poor	Poor	Very poor	Don't know	Response Count
Planning and preparing for evaluation	19	25	2	0	0	1	47
Committing RDA resources and support for evaluation	10	27	7	1	0	3	48
Managing evaluation cost-effectively	10	23	10	3	0	2	48
Ensuring evaluation was carried out robustly	19	19	8	2	0	0	48
Ensuring evaluation met the guidelines of the RDA	19	25	2	0	0	2	48
Ensuring evaluation met the guidelines of BIS	14	20	6	1	0	7	48
Ensuring evaluation met the guidelines of the EU	18	17	6	0	0	7	48
Ensuring transparency in the conduct of evaluation	11	30	4	1	0	2	48
Ensuring independence in the conduct of evaluation	16	26	2	1	0	3	48
Ensuring inclusiveness in the conduct of evaluation	7	29	6	2	0	4	48
Disseminating the results of evaluation within the RDA	1	15	15	12	2	3	48
Disseminating the results of evaluation to external audiences	1	9	20	8	3	7	48
Ensuring learning and development from evaluation findings	1	13	18	12	0	4	48

Table 8: The effectiveness of RDA evaluative research processes – YF sub-population (%)

Question 13(a): How would you rate the evaluation team processes at the RDAs?	YF sub-population (%)						Response Count
	Very good	Good	Neither good nor poor	Poor	Very poor	Don't know	
Planning and preparing for evaluation	40%	53%	4%	0%	0%	2%	47
Committing RDA resources and support for evaluation	21%	56%	15%	2%	0%	6%	48
Managing evaluation cost-effectively	21%	48%	21%	6%	0%	4%	48
Ensuring evaluation was carried out robustly	40%	40%	17%	4%	0%	0%	48
Ensuring evaluation met the guidelines of the RDA	40%	52%	4%	0%	0%	4%	48
Ensuring evaluation met the guidelines of BIS	29%	42%	13%	2%	0%	15%	48
Ensuring evaluation met the guidelines of the EU	38%	35%	13%	0%	0%	15%	48
Ensuring transparency in the conduct of evaluation	23%	63%	8%	2%	0%	4%	48
Ensuring independence in the conduct of evaluation	33%	54%	4%	2%	0%	6%	48
Ensuring inclusiveness in the conduct of evaluation	15%	60%	13%	4%	0%	8%	48
Disseminating the results of evaluation within the RDA	2%	31%	31%	25%	4%	6%	48
Disseminating the results of evaluation to external audiences	2%	19%	42%	17%	6%	15%	48
Ensuring learning and development from evaluation findings	2%	27%	38%	25%	0%	8%	48

Table 9: The effectiveness of RDA evaluative research processes – total population¹¹⁸

Question 17(a): What strategic, operational and relational challenges do you think the evaluation team faced?	Yes	No	Don't know	Response Count
Strategic: Role of evaluation and evaluation processes were not clearly defined within RDA strategic documents (RES etc)	46	26	9	81
Strategic: Changing guidelines and expectations from BIS	57	9	15	81
Strategic: Changing guidelines and expectations from senior management within the RDA	44	19	17	80
Operational: Lack of budget for evaluations	28	43	10	81
Operational: Lack of budget for evaluation team staff	31	25	24	80
Operational: Lack of budget for evaluation team training and development	13	29	39	81
Operational: Lack of budget for dissemination activities	36	22	23	81
Relational: Lack of senior management support for evaluation	42	27	12	81
Relational: Lack of project manager support for evaluation	42	30	9	81
Relational: Lack of external partner's support for evaluation	28	26	26	80

¹¹⁸ For reasons of clarity, in the narrative agreement is denoted by 'Yes' responses.

Table 10: The effectiveness of RDA evaluative research processes – total population (%)

Question 17(a): What strategic, operational and relational challenges do you think the evaluation team faced?	Yes	No	Don't know	Response Count
Strategic: Role of evaluation and evaluation processes were not clearly defined within RDA strategic documents (RES etc)	57%	32%	11%	81
Strategic: Changing guidelines and expectations from BIS	70%	11%	19%	81
Strategic: Changing guidelines and expectations from senior management within the RDA	55%	24%	21%	80
Operational: Lack of budget for evaluations	35%	53%	12%	81
Operational: Lack of budget for evaluation team staff	39%	31%	30%	80
Operational: Lack of budget for evaluation team training and development	16%	36%	48%	81
Operational: Lack of budget for dissemination activities	44%	27%	28%	81
Relational: Lack of senior management support for evaluation	52%	33%	15%	81
Relational: Lack of project manager support for evaluation	52%	37%	11%	81
Relational: Lack of external partner's support for evaluation	35%	33%	33%	80

Table 11: The effectiveness of RDA evaluative research processes –YF sub-population

Question 17(a): What strategic, operational and relational challenges do you think the evaluation team faced?	Yes	No	Don't know	Response Count
Strategic: Role of evaluation and evaluation processes were not clearly defined within RDA strategic documents (RES etc)	30	14	4	48
Strategic: Changing guidelines and expectations from BIS	32	2	14	48
Strategic: Changing guidelines and expectations from senior management within the RDA	31	8	8	47
Operational: Lack of budget for evaluations	21	20	7	48
Operational: Lack of budget for evaluation team staff	19	12	16	47
Operational: Lack of budget for evaluation team training and development	9	14	25	48
Operational: Lack of budget for dissemination activities	24	10	14	48
Relational: Lack of senior management support for evaluation	30	12	6	48
Relational: Lack of project manager support for evaluation	27	15	6	48
Relational: Lack of external partner's support for evaluation	19	13	15	47

Table 12: The effectiveness of RDA evaluative research processes –YF sub-population (%)

Question 17(a): What strategic, operational and relational challenges do you think the evaluation team faced?	Yes	No	Don't know	Response Count
Strategic: Role of evaluation and evaluation processes were not clearly defined within RDA strategic documents (RES etc)	63%	29%	8%	48
Strategic: Changing guidelines and expectations from BIS	67%	4%	29%	48
Strategic: Changing guidelines and expectations from senior management within the RDA	66%	17%	17%	47
Operational: Lack of budget for evaluations	44%	42%	15%	48
Operational: Lack of budget for evaluation team staff	40%	26%	34%	47
Operational: Lack of budget for evaluation team training and development	19%	29%	52%	48
Operational: Lack of budget for dissemination activities	50%	21%	29%	48
Relational: Lack of senior management support for evaluation	63%	25%	13%	48
Relational: Lack of project manager support for evaluation	56%	31%	13%	48
Relational: Lack of external partner's support for evaluation	40%	28%	32%	47

Table 13: The influence of evaluation on specific strategic decision making processes

Question 20(a): Please indicate your level of agreement or disagreement with the following statements concerning the role of evaluation to inform strategic decision making within the RDA.	Strongly agree	Agree	Neither agree or disagree	Disagree	Strongly disagree	Don't know	Response Count
Evaluation evidence fed into the Corporate Plan	1	6	0	1	0	0	8
Evaluation evidence fed into the RES	0	5	2	1	0	0	8
Evaluation evidence fed into the Integrated Regional Strategy	0	3	3	2	0	0	8
Evaluation evidence fed into the ERDF Programme strategy	0	4	3	0	0	1	8
Evaluation evidence fed into decision making at Board meetings	0	3	2	1	0	2	8
Evaluation evidence fed into decision making at Executive meetings	0	4	0	3	0	1	8
Evaluation evidence informed the investment prioritisation exercise due to the Government's budget cuts	0	3	1	3	1	0	8
Evaluation evidence did not inform strategic decision making	0	2	1	5	0	0	8

Appendix 6: Qualitative survey data tables

Table 1: Survey qualitative data tables

Theme	Sub-theme	Quotation	Survey question that generated the quotation ¹¹⁹	Quotation used in thesis narrative?
The principles of regional policy making and the role of RDA evaluation	Using evaluation evidence to promote accountability	There is an obligation on all investors to understand what they have bought and particularly for public sector to be able to account for decisions made with tax payer's money. (External evaluator [75])	51(b): In your opinion do you think conducting evaluations is good value for money? Can you please explain your response?	Y
		Government got into the habit of giving, then taking, responsibilities for too much that was simply being a funding agency and not really a part of the core mission. (External evaluator [76])	27(b): How important were each of the following challenges for the RDA when implementing the budget cuts in 2010? Other (please specify) / further comments	Y
		Directors theoretically buying in to evaluation but prioritising it with senior management only when it was good news or a BIS requirement. (RDA evaluation officer [51])	17(b): What strategic, operational and relational challenges do you think the evaluation team faced? Other (please specify) / further comments	N
		Other RDAs really didn't want objective	17(b): What strategic,	Y

¹¹⁹ See Appendix 2 for survey items and respondents.

		evaluations, just good news (External evaluator [71])	operational and relational challenges do you think the evaluation team faced? Other (please specify) / further comments	
Using evaluation evidence to promote effectiveness		It provides compelling independent evidence of what does and does not work and should improve future delivery. (RDA economic appraisal officer [51]).	51(b): In your opinion do you think conducting evaluations is good value for money? Can you please explain your response?	Y
		It was felt that both project and programme evaluations were important for learning and continuous improvement. (RDA evaluation officer [51])	45(b): Was there a focus on project evaluation or programme level evaluation? What were the reasons for this?	Y
		Yes, wholly agree that public money should be spent on evaluation to improve future performance and impact. Some but not all evaluations achieve that. (RDA strategy officer [64])	51(b): In your opinion do you think conducting evaluations is good value for money? Can you please explain your response?	Y
The uptake and use of evaluation evidence within the RDAs: Strategic decision making	Complexity of using evaluation evidence for strategic decision making	The diverse nature of projects run by the organisation meant that there would never be a way of consistently appraising or evaluating every project and therefore a "gut instinct" approach to project selection was more appropriate and practical. (RDA economic appraisal officer [10])	66: Please use the space below for any additional comments.	Y
		Few projects seemed to be appraised properly.	30: How do you think economic	N

		Given the wide range of projects - urban realm, transport, business development etc this is somewhat inevitable, as the skills to appraise such a wide range of projects were lacking internally (RDA economic appraisal officer [10])	appraisal could/should have been conducted at the RDA?	
		Often spend was the main focus however. Getting money out the door, especially if close to year end, or reigning budgets in if money was tight (particularly towards the end of the life of the RDA). (RDA project delivery/performance officer [12])	42: How do you think monitoring data on spend, outputs and beneficiary data should have been used to inform project delivery at the RDA?	Y
	Unsystematic evaluation evidence use	We didn't really get evaluations to feed into internal and external strategy development work in terms of what works best and cost effectively at delivering outcomes RDA strategy officer [64])	15: What do you think were the main successes and failures of the evaluation team?	Y
		Haphazard investment, politicalisation and end of RDA (RDA project delivery/performance officer [29])	18: What strategic, operational and relational challenges do you think the evaluation team faced? Thinking about your responses to the question above, how do you think the consequences of these challenges played out?	Y
	Role of other factors besides evidence	Ad hoc, driven by Government, local political pressure, internal personalities and response to economic events (RDA strategy officer [60])	23: In general, how was investment prioritisation undertaken at the RDA?	Y
		It could have informed it much more if properly managed - whether it should have done is	21: How do you think evaluation evidence	Y

	Evidence 'cherry picked' to support decisions	another question, were the RDAs there to make strategic decisions or implement policy that had some democratic accountability regardless of what the evidence says, not sure. (External evaluator [74])	could/should have informed strategic decision making at the RDA?	
		It was unclear whether evaluation work was being used to justify projects or guide project development. (RDA economic appraisal officer [10])	14: How would you rate the evaluation team processes at the RDAs? Thinking about your answers above, how do you think evaluation team processes could have been improved and why?	Y
		There are exceptions to this but normally they occurred when a project had to go to central Government and then the agency took it seriously (though it was still an exercise in justifying something that others had already decided should happen). (RDA strategy officer [63])	30: How do you think economic appraisal could/should have been conducted at the RDA?	Y
		Evaluation evidence was presented but I am not sure how much it actually influenced the decisions made. Evaluation evidence was perhaps just used to justify the decisions made. (RDA evaluation officer [53])	18: What strategic, operational and relational challenges do you think the evaluation team faced? Thinking about your responses to the question above, how do you think the consequences of	Y

			these challenges played out?	
		Much of the evaluation completed to 2010 had been skewed towards justifying the RDAs. (External evaluator [73])	How do you think evaluation evidence could/should have informed strategic decision making at the RDA?	Y
		If anything the use of evaluation by the senior team was more an exercise in politics (using evidence to talk up the RDA or seeking to undermine it where it provided a more negative conclusion). (RDA strategy officer [63])	18: What strategic, operational and relational challenges do you think the evaluation team faced? Thinking about your responses to the question above, how do you think the consequences of these challenges played out?	Y
		On occasion [appraisal] was no doubt informed by evaluation too but often this was after the decision was made (i.e. it formed part of the political rationale - find some evidence that supports what you want to do and write it in the business case). (RDA strategy officer [63])	23: In general, how was investment prioritisation undertaken at the RDA?	Y
		The maintenance of politically driven investment schemes with little evidence probably doomed the RDAs. (Central Government officer [82])	21: How do you think evaluation evidence could/should have informed strategic decision making at the RDA?	Y
	Need for a more systematic approach to evidence use	Evaluation evidence should have informed decision making and strategic planning at all levels within the RDA, for example: - Regional Economic Strategy - Investment allocation	21: How do you think evaluation evidence could/should have informed strategic decision making at	Y

	appropriate and practical. (RDA economic appraisal officer [10])			
Unsystematic and cherry picked use of evaluation evidence	There was not genuine buy-in by the Board and as such they did not use evaluation to shape strategy or investment decisions at the time to a sufficient degree (RDA evaluation officer [46])	17(b): What strategic, operational and relational challenges do you think the evaluation team faced? Other (please specify) / further comments	N	
	It should have been undertaken as part of a genuine business case for investment and not as a tool to justify a decision already made. (RDA strategy officer [63])	30: How do you think economic appraisal could/should have been conducted at the RDA?	Y	
The role of benchmarks	[The RDA] never wanted to implement benchmarks as all projects are different. However, this would have been useful in appraisal, and could have been used as a guideline to strive for higher VFM. (RDA project delivery/performance officer [15])	42: How do you think monitoring data on spend, outputs and beneficiary data should have been used to inform project delivery at the RDA?	Y	
	Individual RDAs had a strong evidence based approach and used a basket of ready reckoners and associated models to demonstrate highest impact. Others however made it up as they went. (Central Government officer [82])	33(b): Did evaluation evidence inform economic appraisal at the RDA? If yes, how?	Y	
	Too many academic arguments about who was right and who was wrong about the use of these and a reluctance to start using them because they were a "blunt tool". However there were never really any alternatives put forward. [The	31(b): Were benchmarks (regional or national) used to inform economic appraisal? If not, why not?	Y	

	RDA] took a position whereby we wouldn't use them as a result - other RDAs took the opposite view. (RDA economic appraisal officer [5])		
Issues of generalisability of evaluation evidence	<p>Evaluation evidence was often cited in investment appraisals often more as a placeholder rather than as a serious consideration for projects. The reason is twofold - a - lessons learned is frequently an afterthought and b - actually the pertinence of previous lessons to new project types is not always apparent. It is frequently too easy to generalise the learning points rather than to make them valid for a new activity type. (RDA evaluation officer [53])</p> <p>The evaluation team's input to the appraisal process involved highlighting and applying evaluation findings of previous projects to current projects. This was often difficult, however, as evaluation findings were often quite specific and rarely fully applicable to other projects. (RDA economic appraisal officer [3])</p>	<p>12(b): Please indicate your level of agreement or disagreement with the following statements concerning the role of evaluation within the RDA. Other (please specify) / further comments</p> <p>33(b): Did evaluation evidence inform economic appraisal at the RDA? If yes, how?</p>	Y
Role of other factors besides evidence	Essentially the organisations spending processes did not require the same scrutiny as national projects and I think decision makers were more influenced by political factors in project choice than value for money indicators. (Unless they had to apply for Treasury money in which case it was seen as an obstacle to overcome). (RDA	9(b): Did evaluation evidence inform economic appraisal at the RDA? If yes, how?	Y

	Evaluation as an add on or tick box exercise	economic appraisal officer [10])	Project Managers often saw evaluation as an add-on and came to the evaluation team late to plan or undertake an evaluation. (RDA evaluation officer [56])	17(b): What strategic, operational and relational challenges do you think the evaluation team faced? Other (please specify) / further comments	N
			Evaluation was seen as an 'add on', and not something that was fully part of a project. (RDA project delivery/performance officer [42])	18: What strategic, operational and relational challenges do you think the evaluation team faced? Thinking about your responses to the question above, how do you think the consequences of these challenges played out?	N
			I believe evaluations were a tick box exercise rather than adding any great value. (RDA project delivery/performance officer [27])	18: What strategic, operational and relational challenges do you think the evaluation team faced? Thinking about your responses to the question above, how do you think the consequences of these challenges played out?	N
	How evaluation evidence may shape project appraisal within the wider process		Evaluations should have fed into a framework/matrix which could have been used to influence the appraisal process to a greater extent. (RDA project delivery/performance officer [28])	37: How do you think evaluation could/should have informed project delivery at the RDA?	Y

The uptake and use of evaluation evidence within the RDAs: Delivery	Influence of evaluation on delivery linked to the level of engagement	I've put that project manager support was not a challenge - on the whole, project managers were very supportive, once they realised what was involved and how valuable the findings were but there was a small number of project managers who remained to be convinced. (RDA evaluation officer [58])	17(b): What strategic, operational and relational challenges do you think the evaluation team faced? Other (please specify) / further comments	Y
	Evaluation used for project level learning	Interim project evaluations were very successful at informing the future delivery of projects. In some cases consultants went to the trouble of writing an implementation plan for the recommendations of the evaluation after discussion with the delivery partners and YF project manager. (RDA evaluation officer [56]) I think the evaluation process, while by no means perfect, had a lot going for it. It made project managers think about what their projects were meant to achieve and whether they achieved them. (RDA project delivery/performance officer [39])	36(b): Did evaluation inform project delivery at the RDA? If yes, how?	Y
		I'm never certain how much of the very good learning in the detailed final reports went on to impact future delivery at an operational level, but I'm certain it did in some way impact the way people approached their projects. Very difficult to qualify though. (RDA project delivery/performance officer [12])	36(b): Did evaluation inform project delivery at the RDA? If yes, how?	Y
	Lack of a co-ordinated	I think the individual project interim evaluations	37: How do you think	Y

	learning process	were successful at what they did, but this shared learning should have been expanded out to a similar group of projects to share best practice. (RDA evaluation officer [56])	evaluation could/should have informed project delivery at the RDA?	
The generation of RDA evaluation and the role of 'knowledge brokers'	Strengthening of RDA evaluation practice over time	The main success was to get so many evaluations completed to reasonable standards in a short period of time. The failure was in not having established evaluation as an integral part of the decision making process. (External evaluator [67])	15: What do you think were the main successes and failures of the evaluation team?	Y
		There is little doubt in my mind that evaluation practice moved on significantly over this period. (RDA evaluation officer [53])	48(b): Did the introduction of the Impact Evaluation Framework (IEF) have the following implications? Other (please specify) / further comments	Y
		In general the processes worked well - they improved and became more consistent over time as RDAs became more experienced. (External evaluator [78])	14: How would you rate the evaluation team processes at the RDAs? Thinking about your answers above, how do you think evaluation team processes could have been improved and why?	Y
	Knowledge brokering role provided by evaluation personnel	The team were always approachable, knowledgeable, constructive, helpful and regarded as experts in their field across the organisation. (RDA economic appraisal officer)	15: What do you think were the main successes and failures of the evaluation team?	Y

The communication of RDA evaluation	The need for innovative and targeted dissemination methods	[9]) The team had an excellent reputation both inside and outside the agency for skills, knowledge and effectiveness. An approachable team with the ability to communicate complex and often sensitive evaluation findings in a manner appropriate to the audience. (RDA strategy officer [59])	15: What do you think were the main successes and failures of the evaluation team?	Y
		A great deal of time and energy was invested in the evaluation process, through procuring the most suitably qualified evaluation consultants for the project being evaluated, to setting up robust steering groups to direct the evaluation process, and constructive final meetings to present/clarify results/findings. (RDA Delivery/performance officer [12])	36(b): Did evaluation inform project delivery at the RDA? If yes, how?	Y
		Studies were often highly technical and the varied nature of RDA staff roles meant that key conclusions need to be presented simply so that lessons could be learnt without the technical detail. (RDA economic appraisal officer [10])	14: How would you rate the evaluation team processes at the RDAs? Thinking about your answers above, how do you think evaluation team processes could have been improved and why?	Y
		Final reports are rarely read in full by anyone other than the client project manager yet consultants spend days and weeks writing them. One page summaries, e-shots and newsletter	51(c): In your opinion do you think conducting evaluations is good value for money? Thinking about your answer	Y

		summaries should be used much more to stimulate interest. Obviously a final report is obligatory, but it should be accompanied by more user-friendly outputs to prevent the evaluation dying a death in someone's inbox. (External evaluator [70])	above, how do you think value for money can be improved?	
	Need for greater dissemination	We should have recognised that dissemination of findings and the application of findings to action was about 50% of what we should have been doing. (RDA evaluation officer [56])	14: How would you rate the evaluation team processes at the RDAs? Thinking about your answers above, how do you think evaluation team processes could have been improved and why?	Y
		I would say that we had a lot of work to do on dissemination - we could have contributed to academic journals, put articles into trade journals such as R&R, Building, Estates Gazette, we could have spoken at conferences, we could have led on strategy development - but we failed.	13(b): How would you rate the evaluation team processes at the RDAs? Other (please specify) / further comments	Y
	Lack of analysing and synthesising evaluation evidence	There was a lack of analysis of evaluation and therefore dissemination and learning. (RDA project delivery/performance officer [19])	14: How would you rate the evaluation team processes at the RDAs? Thinking about your answers above, how do you think evaluation team processes could have been improved and	Y

		<p>I'm sure there was lots of useful information contained in the various reports commissioned, but I'm not certain whether the overall learning from all this was combined in a meaningful and concise way so it could influence new projects being developed. (RDA project delivery/performance officer [12])</p>	<p>why?</p> <p>14: How would you rate the evaluation team processes at the RDAs?</p> <p>Thinking about your answers above, how do you think evaluation team processes could have been improved and why?</p>	<p>Y</p>
Displacement of dissemination activities	<p>BIS expectations were a major challenge - resources had to be refocused on producing IEF compliant evaluations rather than evaluations that were useful for the organisation. (RDA evaluation officer [54])</p>	<p>18: What strategic, operational and relational challenges do you think the evaluation team faced?</p> <p>Thinking about your responses to the question above, how do you think the consequences of these challenges played out?</p>	<p>Y</p>	
	<p>The evaluation team had limited resources which were misdirected to achieving compliance with experimental frameworks from both BIS and the EU. These frameworks also demanded a minimum level of RDA spend to be covered by the evaluations. This led to little resources spent on analysing, disseminating and using the findings from the research. We were in effect reacting to the targets set by BIS rather than gaining information that was useful to the region. (RDA evaluation officer [56])</p>	<p>18: What strategic, operational and relational challenges do you think the evaluation team faced?</p> <p>Thinking about your responses to the question above, how do you think the consequences of these challenges played out?</p>	<p>Y</p>	

Barriers and facilitators: a regulatory framework	Evaluation had not been planned for and resourced from the outset	The evaluation team was not established for far too long and should have received more thought in the early stages of the RDA	15: What do you think were the main successes and failures of the evaluation team?	Y
	Evaluation not 'embedded'	The IEF requirements provided the necessary focus internally to undertake evaluation, even though the IEF approach had limitations in assessing overall impact by drilling this down to cost per job. (RDA evaluation officer [49])	13(b): How would you rate the evaluation team processes at the RDAs? Other (please specify) / further comments	Y
		A higher profile and more 'authority' would have helped. (RDA Strategy officer [62])	14: How would you rate the evaluation team processes at the RDAs? Thinking about your answers above, how do you think evaluation team processes could have been improved and why?	Y
		The work of the [evaluation] teams was largely prompted by the requirements of the BIS/PwC work and hence the evaluations were not as well embedded in the decision making procedures of the RDAs as they should have been. (External evaluator [67])	14: How would you rate the evaluation team processes at the RDAs? Thinking about your answers above, how do you think evaluation team processes could have been improved and why?	Y
	Ad hoc demands from central government	A key operational challenge for the evaluation team was to disseminate valuable data that had real meaning. Frequently our evaluation team	17(b): What strategic, operational and relational challenges do you think the	Y

Barriers and facilitators: Political backing		manager would be required to send ad hoc quantitative data to Government and rarely was [he/she] able/allowed to provide a brief description of what the data actually meant and its limitations, i.e. place it in a qualitative setting. (RDA evaluation officer [56])	evaluation team faced? Other (please specify) / further comments	
	Formal acknowledgement of using evaluation evidence	I think the lack of recognition of the importance of evaluation in both strategic and corporate documentation prevented the full benefits from being realised. This meant that projects were not as efficient or effective as they could be and we did not have the robust assessment of our impact that we should have done.	18: What strategic, operational and relational challenges do you think the evaluation team faced? Thinking about your responses to the question above, how do you think the consequences of these challenges played out?	Y
	Senior manager support for evaluation	Senior exec level commitment massively eased the challenges of persuading operational teams to engage positively with evaluation. (RDA evaluation officer [55])	18: What strategic, operational and relational challenges do you think the evaluation team faced? Thinking about your responses to the question above, how do you think the consequences of these challenges played out?	Y
		The senior team played lip service to evaluation and whilst evaluation could have been viewed as essential to the evidence based policy making described in strategy documents the reality was at odds with this. (RDA strategy officer [63])	18: What strategic, operational and relational challenges do you think the evaluation team faced? Thinking about your responses to the question above, how do	Y

			you think the consequences of these challenges played out?	
		I was never really sure what senior management at [the RDA] thought about evaluation. Perhaps if they had been more vocal in their support both to [RDA] staff and our project delivery partners then it may have smoothed some problems over. (RDA project delivery/performance officer [12])	18: What strategic, operational and relational challenges do you think the evaluation team faced? Thinking about your responses to the question above, how do you think the consequences of these challenges played out?	Y
	Project manager support for evaluation	Project managers sometimes seemed to view evaluation as being rather burdensome and of little relevance to their work. (RDA strategy officer [59])	18: What strategic, operational and relational challenges do you think the evaluation team faced? Thinking about your responses to the question above, how do you think the consequences of these challenges played out?	Y
		Project Managers often saw evaluation as an add-on and came to the evaluation team late to plan or undertake an evaluation. Therefore it was very difficult for the team to effectively plan its own time accordingly. (RDA evaluation officer [56])	17(b): What strategic, operational and relational challenges do you think the evaluation team faced? Other (please specify) / further comments	Y
	Delivery partner support for evaluation	No one (project managers or partners) like evaluation and as "a necessary evil" it does not get the priority it could. (RDA project delivery/performance officer [18])	17(b): What strategic, operational and relational challenges do you think the evaluation team faced?	Y

			Other (please specify) / further comments	
		Sometimes it was a nightmare. Contractors didn't really understand why external consultants were brought in to evaluate and they were really wary of them. (RDA project delivery/performance officer [12])	18: What strategic, operational and relational challenges do you think the evaluation team faced? Thinking about your responses to the question above, how do you think the consequences of these challenges played out?	Y
Evaluation linked to investment decisions		When there was no obvious or direct link between the evaluation and a specific element of further investment, it was very hard to engage with partners. This meant that most evaluations lacked support and lacked interest, meaning that they were ultimately poor value for money. This is a pity, because the quality of the evaluations themselves was high. (RDA economic appraisal officer [3])	18: What strategic, operational and relational challenges do you think the evaluation team faced? Thinking about your responses to the question above, how do you think the consequences of these challenges played out?	Y
Lack of an evaluation culture		Main failures: Inability to raise the profile of evaluation with YF, inability to genuinely communicate evaluation findings in a format that was accessible to potential users, too academic focused and not enough focus on the real end point of evaluation which to me is its practical application to improve project delivery and investment prioritisation. This being said the team's ability on this was significantly hampered	15: What do you think were the main successes and failures of the evaluation team?	Y

		by the general ambivalence to evaluation shown by the Chief Exec and Directors and strategy team (RDA strategy officer [63])			
		The dissemination to external (and even internal) audiences was on a lower than ideal level, not because we were protective / un-transparent, but because there was not that much interest in what our findings were! (RDA economic appraisal officer [3])	13(b): How would you rate the evaluation team processes at the RDAs? Other (please specify) / further comments	Y	
		The culture within [the RDA] was not conducive to using evaluation evidence. (RDA strategy officer [63]).	13(b): How would you rate the evaluation team processes at the RDAs? Other (please specify) / further comments	Y	
		Championing from the top We did not have anyone outside of the evaluation team championing or really understanding evaluation, so our voice within [the RDA] was not very loud. (RDA evaluation officer [56])	18: What strategic, operational and relational challenges do you think the evaluation team faced? Thinking about your responses to the question above, how do you think the consequences of these challenges played out?	N	
		Think because evaluation was not used as a tool it was perceived as an add on rather than being valuable always. There should have been a senior manager to champion this. (RDA evaluation officer [57])	18: What strategic, operational and relational challenges do you think the evaluation team faced? Thinking about your responses to the question above, how do	N	

Barriers and facilitators: Analytical skill	Need for capacity building	Economic appraisal suffered from the same issue as evaluation. It was not central to decision making largely because it was not understood (RDA strategy officer [63]) Support for, and a broad understanding of, evaluation is crucial at the top of the organisation to create a culture of evaluation. (RDA evaluation officer [49])	you think the consequences of these challenges played out? 28: How was the economic appraisal of potential projects and programmes carried out at the RDA? 14: How would you rate the evaluation team processes at the RDAs? Thinking about your answers above, how do you think evaluation team processes could have been improved and why?	Y
Barriers and facilitators: Data quality and availability	Lack of a common evaluation framework	Evaluation should have been considered at the outset - establishing a framework to inform monitoring approaches thereby enabling a robust evaluation to be undertaken. Too often it seemed to be an afterthought with little evidence efficiently obtainable to inform a robust evaluation. (External evaluator [79]) Evaluations would have been more robust and informative had they been carried out systematically and comprehensively over a longer period of time and as an integral part of the investment cycle. (External evaluator [67]) There could have been a more standardised	46: How do you think projects and programmes could/should have been evaluated at the RDA(s)? 51(c): In your opinion do you think conducting evaluations is good value for money? Thinking about your answer above, how do you think value for money can be improved? 14: How would you rate the	Y

	approach to reporting and methods within RDAs to ensure comparability across findings. (External evaluator [75])	evaluation team processes at the RDAs? Thinking about your answers above, how do you think evaluation team processes could have been improved and why?	
	Within a more universal framework and with a common series of outcomes/outputs in mind. The lack of the latter proved fatal. (Central Government officer [82])	46: How do you think projects and programmes could/should have been evaluated at the RDA(s)?	Y
	I think working out a consistent model for what should be measured and how the information should be collected would have been helpful. (RDA evaluation officer [57])	46: How do you think projects and programmes could/should have been evaluated at the RDA(s)?	Y
Need for holistic approach to monitoring and evaluation	Should be a dynamic process of collaboration to not ensure delivery but to measure impact. (RDA project delivery/performance officer [38])	40: How do you think projects could/should have been monitored at the RDA?	Y
	Formal evaluation and quantitative monitoring should have been more closely aligned. There was a complete disconnect between the very contractual/quantitative monitoring of the project carried out through various PMF report templates primarily focusing on, spend, outputs, milestones etc, and the qualitative learning coming out of the evaluation process. (RDA project delivery/performance officer [12])	40: How do you think projects could/should have been monitored at the RDA?	Y
Need for contractual	For the most part, primary data was collected at	56: What primary data	Y

	agreement of beneficiary data	the end of the project and often businesses did not recall in detail the assistance provided - it would have been better to carry out on-going monitoring for evaluation. (RDA evaluation officer [48])	collection and analysis methods were used for evaluations at the RDA? How do you think the use of primary data could have been improved?	
	Resulting issues with both the quality and availability of monitoring data and evaluation data	If the quality and quantity of both evaluation AND monitoring data been improved, it might have been possible to set up systems to make performance (including return on investment and social and environmental performance) within and across projects more visible to the strategic decision makers. (RDA evaluation officer [58])	21: How do you think evaluation evidence could/should have informed strategic decision making at the RDA?	Y
		Many RDA evaluations, particularly quantitative assessment are of poor quality or of variable quality, which makes it difficult to synthesise. (RDA evaluation officer [46])	53(b): What do you think are the main challenges for synthesising the current evaluation evidence base to inform decision making for the LEPs and the Regional Growth Fund? Other (please specify) / further comments	Y
	Inappropriate to manipulate data to try and make it 'fit' into decision-relevant formats retrospectively	We were starting to use evaluation evidence as a benchmark to see if the project was expecting to produce a return on investment comparable with previous similar projects. There were dangers there though because the benchmark	33(b): Did evaluation evidence inform economic appraisal at the RDA? If yes, how?	Y

		data were not derived from a process that was designed to provide benchmark data. (RDA evaluation officer [58])		
		Decision makers like evaluation to provide simple 'answers' - such as return on investment measures. In reality it is dangerous to base decisions on these - because of methodological limitations (e.g. not costing certain sorts of benefits and inconsistencies in treatment) and changing contexts. (RDA evaluation officer [54])	21: How do you think evaluation evidence could/should have informed strategic decision making at the RDA?	Y
	Lack of policy relevance of evaluation evidence	Main failures: Inability to raise the profile of evaluation with [the RDA], inability to genuinely communicate evaluation findings in a format that was accessible to potential users, too academic focused and not enough focus on the real end point of evaluation which to me is its practical application to improve project delivery and investment prioritisation. (RDA strategy officer [63])	15: What do you think were the main successes and failures of the evaluation team?	Y
	IEF methodology mechanistic	The IEF requirements provided the necessary focus internally to undertake evaluation, even though the IEF approach had limitations in assessing overall impact by drilling this down to cost per job. (RDA evaluation officer [49])	13(b): How would you rate the evaluation team processes at the RDAs? Other (please specify) / further comments	Y
		One of the weaknesses has been an inability to demonstrate the added value and spillover of regeneration and social investment. Anything that could be done in that field that HMT will	25: How was evaluation evidence used to inform the budget cuts investment prioritisation exercise?	Y

	accept would be invaluable. (Central Government officer [82])		
	IEF is grossly mechanistic in its focus upon all the theoretical steps in the gross-net adjustment, most of which are incapable of empirical research or involve application of standard assumptions. (External evaluator [73])	53(b): What do you think are the main challenges for synthesising the current evaluation evidence base to inform decision making for the LEPs and the Regional Growth Fund? Other (please specify) / further comments	Y
	Context is the big issue - context, mechanism and outcomes are rarely differentiated. (RDA evaluation officer [54])	53(b): What do you think are the main challenges for synthesising the current evaluation evidence base to inform decision making for the LEPs and the Regional Growth Fund? Other (please specify) / further comments	Y
IEF methodology did not produce policy learning	The highly restrictive (and experimental) impact evaluation framework (IEF) set out by BIS meant a great deal of resources were spent collecting data that was not helpful at project level and was neither used at a strategic decision making level. (RDA evaluation officer [56])	12(b): Please indicate your level of agreement or disagreement with the following statements concerning the role of evaluation within the RDA. Other (please specify) / further comments	Y

	More focus on trying to get evaluations and numbers that fitted the narrow definitions adopted by PWC and a complete disinterest in learning from evaluation. (RDA evaluation officer [51])	49(b): Did the PWC exercise to collate evaluation evidence for the 'Impact of RDA spending: national report' for BERR have the following implications? Other (please specify) / further comments	Y
		18: What strategic, operational and relational challenges do you think the evaluation team faced? Thinking about your responses to the question above, how do you think the consequences of these challenges played out?	Y
		51(b): In your opinion do you think conducting evaluations is good value for money? Can you please explain your response?	Y
		51(b): In your opinion do you think conducting evaluations is good value for money? Can you please explain your response?	Y
	RDA evaluation findings and processes were undermined	53(b): What do you think are the main challenges for synthesising the current	Y

		involvement in evaluation over the longer term. (External evaluator [73])	evaluation evidence base to inform decision making for the LEPs and the Regional Growth Fund? Other (please specify) / further comments	
		Evaluation will always be the poor relation to other more pressing needs, particularly given that if a) few believe the findings (RDA impact report) or b) if few take the time to learn the lessons, then it will not be a departmental priority - especially in times of austerity. (RDA evaluation officer [53])	18: What strategic, operational and relational challenges do you think the evaluation team faced? Thinking about your responses to the question above, how do you think the consequences of these challenges played out?	Y
	RDAs failed to utilise evaluation evidence to promote effectiveness	Danger of evaluation being seen as something you do because you have to rather than want to. Also as ever, political imperative for some projects can get in the way of good intentions. Consequences would be poor projects which don't deliver benefits, and not able to demonstrate adequately effectiveness - this has been crucial in the demise of RDAs! (RDA project delivery/performance officer [31]) I feel that an inability to make the case for the RDAs in terms of added value doomed them ultimately. (Central Government officer [82])	18: What strategic, operational and relational challenges do you think the evaluation team faced? Thinking about your responses to the question above, how do you think the consequences of these challenges played out?	Y
			18: What strategic, operational and relational challenges do you think the evaluation team faced? Thinking about your responses	Y

A potential role for knowledge translation tools	Volume of evidence requiring critical appraisal	I think it was, but the information was lost in the waves of constant information fired at project managers. (RDA project delivery/performance officer [21])	to the question above, how do you think the consequences of these challenges played out?	Y	
		The quantity of information will put off many people from even looking at the evaluation reports. (RDA evaluation officer[52])	42: How do you think monitoring data on spend, outputs and beneficiary data should have been used to inform project delivery at the RDA? 53(b): What do you think are the main challenges for synthesising the current evaluation evidence base to inform decision making for the LEPs and the Regional Growth Fund? Other (please specify) / further comments	Y	
	Accessing and navigating evidence	Evaluation reports should be more easily accessible internally and externally, e.g. reports posted on-line. (Delivery/Performance RDA officer [25])	42: How do you think monitoring data on spend, outputs and beneficiary data should have been used to inform project delivery at the RDA?	Y	
		I don't feel that findings from evaluations were freely available in order to help develop projects, especially cross-team and cross-directorate. (RDA project delivery/performance officer [11])	12(b): Please indicate your level of agreement or disagreement with the following statements	Y	

			concerning the role of evaluation within the RDA. Other (please specify) / further comments	
Systematic approach to evidence use	I think having a more systematic approach to the exploitation of findings would have helped. (RDA evaluation officer [58])		63: How were evaluation findings tested to produce recommendations at the RDA? How do you think this could have been improved?	Y
The use of web-based repositories and databases	Mainly around stronger objective guidelines and the assembling of evaluation findings in a way which could be readily accessible i.e. some sort of interrogatable database. (RDA project delivery/performance officer [40])		14: How would you rate the evaluation team processes at the RDAs? Thinking about your answers above, how do you think evaluation team processes could have been improved and why?	Y
	Again, the ability to search and access data/information easily and remotely.		37: How do you think evaluation could/should have informed project delivery at the RDA?	Y

Appendix 7: Survey items and respondents - online workshop

Table 1: Survey items overview – online workshop

Area	No.	Question	Format	Response options
Identifier	1	Name (kept anonymous)	Open-ended	
Scenario 1	2	Rank the projects from 1 to 10 (1 = most favourable project to receive investment, 10 = least favourable project to receive investment).	Matrix of choices (single select per row)	<u>Row choices:</u> <ol style="list-style-type: none"> Digital House Regional events Atown market regeneration Charity support Atown railway station Encouraging enterprise Saving the planet Internship for graduates Innovating together Opportunities for all <u>Column choices:</u> 1-10
	3	Which projects would you take forward (up to the £5m budget)?	Matrix of choices (single select per row)	<u>Row choices:</u> <ol style="list-style-type: none"> Digital House Regional events Atown market regeneration Charity support Atown railway station Encouraging enterprise Saving the planet

				8. Internship for graduates 9. Innovating together 10. Opportunities for all <u>Column choices:</u> Yes No
4	What is the total cost of the projects you have chosen?		Open-ended	
5	In real life how would you have decided which projects should go forward?		Open-ended	
6 (a)	What level of importance did you give the different pieces of information?		Matrix of choices (single select per row)	<u>Row choices:</u> Project name/description Local Authority Forecast investment 2012-2013 Total GVA in £s (for investment 2012-13) Jobs created Businesses created Businesses supported People assisted to gain employment People assisted in skills CO2 reduced (tonnes) Political/pragmatic considerations <u>Column choices:</u> Most important Very important

				Quite important Important Not very important
6 (b)	Please explain your decisions above		Open-ended	
7(a)	What factors do you think are important when coming to a decision?		Matrix of choices (single select per row)	<u>Row choices:</u> Monitoring data Evaluation data Socio-economic data Contract costs data Time pressure Background and experience Personal values Personal expectations Personal interests Policies and procedures Organisational hierarchy Organisational politics The market in which the organisation operates The economy Government legislation Partner organisation's anticipated reaction Public's anticipated reaction <u>Column choices:</u> Most important Very important

				Quite important Important Not very important
	7(b)	Please list any other factors you think are important	Open-ended	
	8(a)	How easy/difficult do you think it would be to rank the projects if:	Matrix of choices (single select per row)	<u>Row choices:</u> The number of projects increased There was more information given for each project A small team needed to come to a decision A large team needed to come to a decision There was a clear strategy i.e. to reduce short term unemployment There was no clear strategy <u>Column choices:</u> Very easy Easy Neither easy or difficult Difficult Very difficult
	8(b)	Please explain your decisions above	Open-ended	
Scenario 2	9	How do you think the process of decision making can be made replicable and transparent?	Open-ended	
	10	How do you think this scenario compares to scenario 1?	Open-ended	
	11(a)	Would you use such a tool for decision making?	Multiple choice (single select)	Definitely would use

				Likely to use Not likely to use Definitely wouldn't use
11(b)	Please explain your decision above		Open-ended	
12(a)	What are the positives of using such decision tools?		Matrix of choices (single select per row)	<u>Row choices:</u> Assists collaborative decision making Uses available data Decision makers can work with analysts to ensure tool is appropriate Enables decisions to be replicable Enables decisions to be transparent Enables total benefits and costs to be calculated Flexibility if strategic/budget decisions change Would enable better outcomes (GVA/outputs etc) <u>Column choices:</u> Strongly agree Agree Neither agree nor disagree Disagree Strongly disagree
12(b)	Other & please explain your decisions above		Open-ended	

	13(a)	What are the negatives of using such decision tools?	Matrix of choices (single select per row)	<p><u>Row choices:</u></p> <p>Poor monitoring data</p> <p>Poor evaluation data</p> <p>Poor socio-economic data</p> <p>Poor contracting cost data</p> <p>Unable to capture all qualitative information</p> <p>Hidden assumptions</p> <p>Difficult to understand the tool</p> <p>Potential mistakes in the programming of the tool</p> <p>Would not enable better outcomes (GVA/outputs etc)</p> <p><u>Column choices:</u></p> <p>Strongly agree</p> <p>Agree</p> <p>Neither agree nor disagree</p> <p>Disagree</p> <p>Strongly disagree</p>
	13(b)	Other & please explain your decisions above	Open-ended	
	14(a)	Would such a decision tool have been useful for prioritising investment:	Matrix of choices (single select per row)	<p><u>Row choices:</u></p> <p>At the beginning of the recession in the RDAs and BIS</p> <p>Due to the austerity cuts/closure of the RDA network in the RDAs and BIS</p> <p>Currently/in the future for the</p>

				<p>Regional Growth Fund Currently/in the future for the Growth and Innovation Fund Currently/in the future for the Local Enterprise Partnerships Currently/in the future for the Local Authorities</p> <p><u>Column choices:</u> Strongly agree Agree Disagree Strongly disagree Don't know</p>
14(b)	Other & please explain your decisions above	Open-ended		
15	From your experience, how do you think decision making processes for prioritising investment can be improved?	Open-ended		
16	From your experience, what do you think is the best way to package evidence to feed through into decision making processes for prioritising investment?	Open-ended		

Appendix 8: Additional screenshots of the online workshop

Figure 1: Introduction web-page screenshot

WORKSHOP: DECISION MAKING

Investment Prioritisation in UK Regional Policy

HOME

MY PAGE

WORKSHOP

FORUM

ADMIN

MY NETWORK

Options

+

Add

All Pages

My Pages

Introduction

Hello and thank you for taking the time to participate in this online workshop.

Prior to the workshop you received a participant information sheet, which can be downloaded again [HERE](#)

What is the purpose of the workshop?

- The purpose of the research is to investigate what evidence is useful and how it can be packaged effectively to inform decision making processes.
- The aim is to explore the ways that policy makers synthesise different types of information when they make decisions and to study the ways that evaluation, monitoring, socio-economic data and other information can be used to inform these decisions.
- As a case study, investment prioritisation decisions within Regional Policy making is the focus of the research.
- The purpose of the online workshop is to gather data on decision making processes and to explore any lessons learned to inform future recommendations.
- There are no right and wrong answers. It is the thought processes that you go through and the usefulness of evidence to assist those decisions which is the focus of the research.

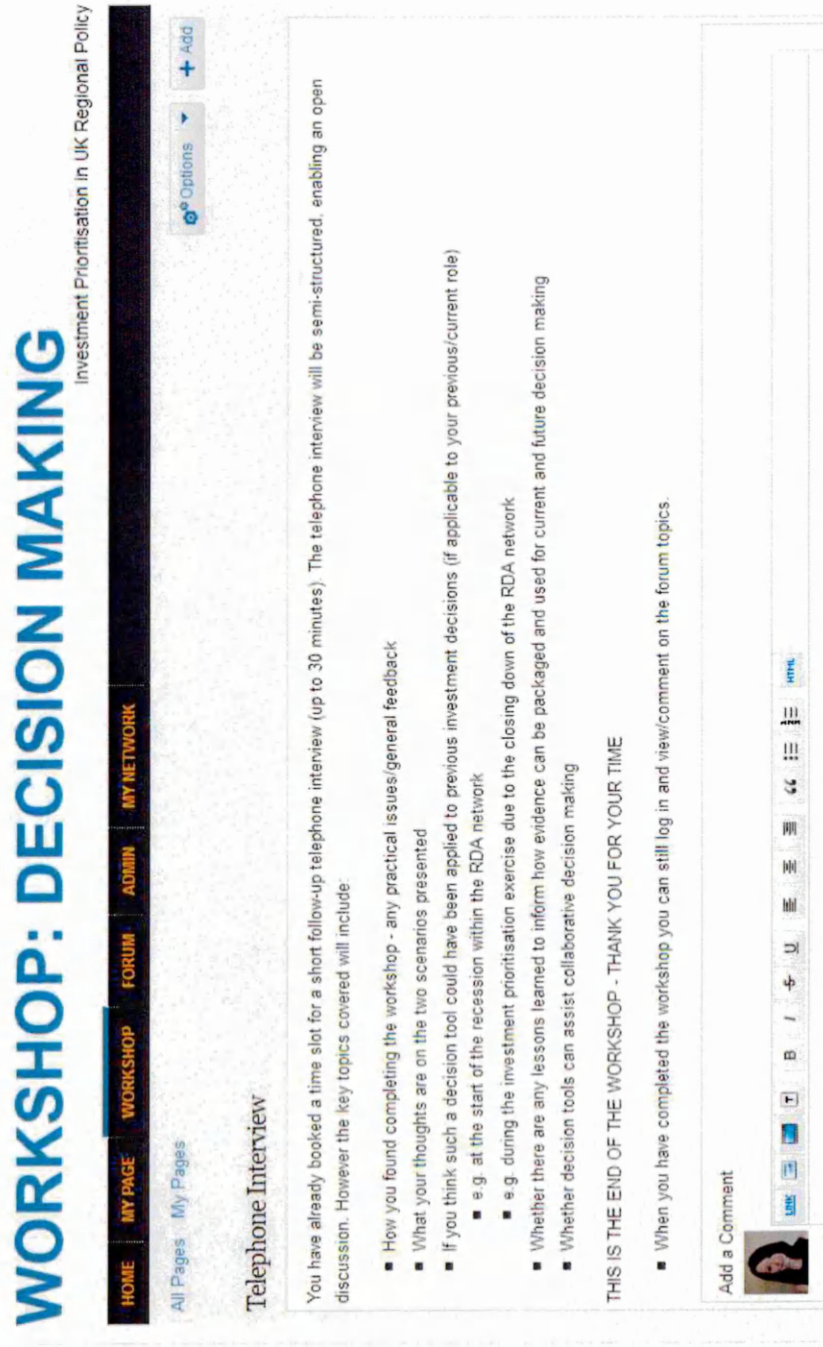
How is the workshop organised?

- You can log in and out of the workshop when you like using your email and password.
- You need to work through the workshop in order. Instructions are given for each part of the workshop.
- You will need to spend approximately 30 minutes on scenario 1 and 20 minutes on scenario 2.
- The workshop will be followed up with a short telephone interview (up to 30 minutes). You need to complete the workshop prior to the telephone interview.
- An overview of what will be discussed in the telephone interview is given at the end of the workshop.

How do I begin?

- Start with scenario 1 and complete both of the scenarios according to the instructions.
- Answer the questions (or start a new topic) on the public forums to discuss the issues with fellow workshop participants.
- Use the private 'chat' function to ask the workshop administrator questions, give feedback or highlight any issues.

Figure 2: Telephone interview web-page screenshot



Appendix 9: Information provided for Scenario 1

Table 1: Information provided for scenario 1

Project Summary			Project Investment	Economic Impact	Expected Outputs					Other information	
Project name	Project description	Local Authority	Forecast investment 2011-2012	Total GVA in £s (for investment 2011-12)	Jobs created	Businesses created	Businesses supported	People assisted to gain employment	People assisted in skills	CO2 reduced (tonnes)	Political/pragmatic considerations
1. Digital House	Building purchased to develop into high quality office accommodation for digital companies	Barnsley District	£150,000	£3,075,000	100	0	0	0	0	0	Took 5 years to acquire working with the local authority. Currently an eyesore in the centre of the town. A lot of negative publicity in the local press about how long it has taken to begin development.
2. Business share network	Funding for business networks to bring companies together to share knowledge and experience with the aim of improving performance	Whole region	£1,500,000	£21,150,000	0	1	30	0	0	0	There have been a couple of high profile business collaborations and innovations due to this network which have been positive publicity. A local MP chairs the board.
3. Rural opportunities	Funding to increase access to transport and training opportunities for people in rural communities targeted at unemployed and unskilled people	Whole region	£80,000	£1,128,000	0	0	0	10	0	0	It is forecast that the level of unemployed are going to rise in rural areas and transport is likely to be cut in budget cuts.
4. Regional events	Funding for regional events to raise the profile of the region and attract inward trade, investment and visitor spend	Whole region	£200,000	£2,820,000	0	0	0	0	0	0	It is difficult to capture the impact of regional events. However they are high profile events.
5. Connecting to London	Funding one additional train per hour between Leeds and London	Leeds District	£2,000,000	£27,200,000	0	0	0	0	0	0	It has taken 3 years to negotiate this deal. There is a large contract set up with large settlement costs if the project is stopped.

6. Centre for sustainability	Research centre to research all aspects of low carbon economy	Whole region	£2,000,000	£17,400,000	100	10	300	0	300	200000	A number of high profile academics are tied into this project. Sustainability is going to be high on the Government's agenda going forward.
7. Regional flood defences	Regional flood defence planning and funding	Whole region	£600,000	£4,980,000	0	0	0	0	0	0	There has been a lot of negative publicity surrounding the recent flooding's. It is possible to leverage in a lot of money from European funds for this project.
8. Business support programme	Business support service across the region (phone/website/courses/access to other support)	Whole region	£13,000,000	£84,500,000	10000	3500	40000	0	0	0	Evaluation reports show that it is important to link businesses to business support providers, but that the impact is difficult to quantify. It is likely the Government will want to reduce these types of projects across the UK.
9. Atown market regeneration	Redevelopment of town market complex into a theatre, café and town council offices	North Lincolnshire	£300,000	£4,230,000	5	5	20	0	0	0	Most of this project has been carried out. This remaining budget is for completion and doing an evaluation to look at the impact of the investment.
10. Charity support	Support, advice and funding for the voluntary and community sector	Whole region	£2,000,000	£19,600,000	0	2	70	0	0	0	The charity sector is going to become more important going forward as public services are cut.
11. Atown railway station	Construction of a new railway station	Rotherham District	£50,000	£705,000	5	2	0	0	0	0	This funding is strategically important to make other partners invest a lot of money into the station.
12. Sports academy	Construction of sports academy to expand the town's leisure and tourism offer	East Riding of Yorkshire	£500,000	£4,900,000	10	2	0	0	0	35	This has received a lot of publicity from high profile sports personalities and MPs. Events have already been planned for the opening and in the future.

13. Broadband access	Broadband provision several years ahead of the commercial market	Whole region	£6,000,000	£58,800,000	40	1	0	0	0	0	This project has received a lot of publicity and is part of the Government's national strategy to have total broadband coverage.
14. Encouraging enterprise	Delivery of a programme in primary and secondary schools to increase the number of young people starting businesses	Whole region	£250,000	£3,525,000	0	0	0	0	3000	0	This project has already been contracted to a delivery company who have hired a team of 30 people.
15. Saving the planet	Raising awareness of sustainability issues to increase business use and investment in low carbon technologies	Whole region	£220,000	£3,102,000	10	4	500	0	0	150000	The government is going to be introducing new laws so that businesses must improve their use of low carbon technologies.
16. Internship for graduates	An employability scheme to assist the rising number of unemployed graduates accessing internship opportunities	Whole region	£250,000	£2,650,000	350	0	800	4000	2000	0	The Government is likely to introduce a new policy to increase the level of internships.
17. Innovating together	Programme aimed at increasing competitiveness of manufacturing businesses through collaboration with universities	Whole region	£200,000	£1,660,000	30	1	5	0	0	0	This programme has been running successfully for a number of years. Many new innovations are in the early stages of development.
18. Opportunities for all	Funding to increase access to transport and training opportunities for people in deprived communities	Whole region	£300,000	£2,490,000	0	0	0	100	50	0	It is forecast that the level of unemployed are going to rise in deprived areas and transport is likely to be cut in budget cuts. There are a number of private funders involved and settlement costs will need to be agreed if the contract is stopped.
19. Women entrepreneurs network	Support for a network of female business owners and leaders	Whole region	£100,000	£1,410,000	0	0	200	0	0	0	The Government wants to showcase projects which support women setting up businesses.
20. Forward together	Business support project to encourage business start-up and growth in Atown	Rotherham District	£300,000	£4,230,000	100	10	50	0	0	0	This town is the most deprived area of the region and there is a lot of pressure from local councillors for this project to go ahead.

Appendix 10: Follow-up telephone interview topic guide

- How you found completing the workshop - any practical issues/general feedback
- What your thoughts are on the two scenarios presented
- If you think such a decision tool could have been applied to previous investment decisions (if applicable to your previous/current role)
 - e.g. at the start of the recession within the RDA network
 - e.g. during the investment prioritisation exercise due to the closing down of the RDA network
- Whether there are any lessons learned to inform how evidence can be packaged and used for current and future decision making
- Whether decision tools can assist collaborative decision making

Appendix 11: Online workshop participant information sheet

Due to your experience and skills you have been selected as a candidate to participate in an important research project focused on evidence use and its role in decision making processes. The research is part of a PhD study being conducted at the Centre for Regional Economic and Social Research based at Sheffield Hallam University and the results of the study are to be published. Before you decide whether or not to take part, it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully.

What is the purpose of the study?

The purpose of the research is to investigate what evidence is useful and how it can be packaged effectively to inform decision making processes. The aim is to explore the ways that policy makers synthesise different types of information when they make decisions and to study the ways that evaluation, monitoring, socio-economic data and other information can be used to inform these decisions. As a case study, investment prioritisation decisions within Regional Policy making is the focus of the research. The purpose of the online workshop is to gather data on decision making processes and to explore any lessons learned to inform future recommendations.

Why have I been invited to participate?

Due to your experience and skills you have been selected as a candidate for this workshop and your contribution will enable a more rounded and balanced picture to be developed. Participants from BIS (Department for Business, Innovation and Skills), and from the former emda (East Midlands Development Agency) and ONE (One North East) have been carefully selected from a range of roles. Candidates have been chosen who were either involved in past decision making processes through the RDA network or who are currently involved in investment decision making processes at BIS.

Do I have to take part?

It is up to you to decide whether or not to take part. If you do decide to take part you need to fill in the consent form below. If you decide to take part you are still free to withdraw at any time and without giving a reason.

What will I have to do if I take part?

You will be sent an email inviting you to join the online workshop on Monday 25th June at 9am. You need to log in as soon as possible and register your details and view the other workshop participants. Then you will be able to log in and out whenever is convenient for you over the course of the week. The workshop will take no more than one hour to complete. In the workshop you will be guided through a scenario focusing on how evidence

can be packaged to feed through into decision making. You will be given a brief introduction and then there will be a short example to work through. There will then be a short introduction video to a decision tool that has been developed to assist decision making, which you will be asked to comment on. There are no right and wrong answers. It is the thought processes that you go through and the usefulness of evidence to assist those decisions which is the focus of the research.

The workshop will be followed up with a short telephone interview (up to 30 minutes) for the researcher to ask participants how they found completing the workshop. Please note that this phone call will be recorded.

How will the information be used?

The information from the workshop and the telephone interviews will be transcribed. The transcript will not be used for any other purpose. The information from the workshop and the telephone interviews will form the basis of the PhD thesis which will be assessed. The transcripts will also be used to write and publish articles in academic journals. You are welcome to see a copy of the transcript for your contribution, and/ or the final thesis and/ or a copy of the articles before they are published.

If I agree to take part what happens to what I say?

All the information you give will be confidential and used for the purposes of this study only. The data will be collected and stored in accordance with the Data Protection Act 1998 and will be disposed of in a secure manner. The information will be used in a way that will not allow you to be identified individually. If you decide to withdraw from the research within 2 weeks of the workshop finishing, all data will be destroyed. After this point data cannot be withdrawn but will always remain anonymised.

The research has the approval of the Faculty Research Ethics Committee and is funded by the Economic and Social Research Council. The student conducting this research project is Jessica Baxendale, supervised by Professor Paul Lawless, Professor Peter Wells and Ian Wilson. Jessica Baxendale was the former Evaluation Analyst at Yorkshire Forward working with Sue Richardson, Bea Jefferson and Lorna Hewish.

What do I do now?

Think about the information above and email: Jessica.K.Baxendale@student.shu.ac.uk if you have any questions or concerns.

If you would like to take part, please fill in the consent form below now. Please note that submitting the online consent form is understood to mean

that informed consent to participate in the study has been given.

THANK YOU FOR YOUR TIME.

Date: 2/5/2012

* 1. What is your full name?

* 2. Please answer the following questions:

	Yes	No
Have you read and understood the information above about this study?	<input type="radio"/>	<input type="radio"/>
Have you been able to ask questions about this study?	<input type="radio"/>	<input type="radio"/>
Have you received enough information about this study?	<input type="radio"/>	<input type="radio"/>
Do you understand that you are free to withdraw from this study at any time?	<input type="radio"/>	<input type="radio"/>
Do you understand that you are free to withdraw from this study without giving a reason for your withdrawal?	<input type="radio"/>	<input type="radio"/>
Your responses will be anonymised before they are analysed: Do you give permission for members of the research team to have access to your anonymised responses?	<input type="radio"/>	<input type="radio"/>
Do you agree to take part in this study?	<input type="radio"/>	<input type="radio"/>

* 3. Please select a time for your telephone interview by copying this link into your browser: <http://www.doodle.com/x6tm4izrpx5kaf5d#table>

Have you successfully booked a time?

Yes

No - please email Jessica.K.Baxendale@student.shu.ac.uk

* 4. What phone number is it best to contact you on?

10

Please print out a copy of this consent form and information sheet for your records.

Contact details:

Jessica Baxendale
Centre for Regional Economic and Social Research
Sheffield Hallam University

City Campus
Howard Street
Sheffield
S1 1WB

Phone X
Mobile X
E-mail X

Appendix 12: Online workshop email communication

INTRODUCTION EMAIL:

Dear [x]

Your contact details have been passed onto me by [emda gatekeeper name/ONE gatekeeper name /BIS gatekeeper name].

Due to your experience and skills gained from your [current/previous] role at [emda/ONE/BIS] you have been selected as a candidate to participate in an important research project focused on evidence use and its role in decision making processes. The research is part of a PhD study being conducted at the Centre for Regional Economic and Social Research based at Sheffield Hallam University and the results of the study are to be published.

Participants from BIS (Department for Business, Innovation and Skills), and from the former emda (East Midlands Development Agency) and ONE (One North East) have been carefully selected from a range of roles. Your contribution will enable a more rounded and balanced picture to be developed and to capture lessons learned.

For more information and to decide whether or not you would like to take part, please now view the participant information sheet: [HERE](#)

If you have any questions or concerns, please email: Jessica.K.Baxendale@student.shu.ac.uk.

Best wishes,

Jess

Jessica Baxendale

PhD Student

Centre for Regional Economic and Social Research
Sheffield Hallam University

City Campus
Howard Street
Sheffield
S1 1WB
Phone: x

AUTOMATIC EMAIL TO BE SENT AT END OF CONSENT FORM

Many thanks for agreeing to participate in the online workshop exploring evidence use and its role in decision making processes.

You will be sent an email inviting you to join the online workshop on Monday 25th June at 9am. You need to log in as soon as possible and register your details and view the other workshop participants. Then you will be able to log in and out whenever is convenient for you.

The workshop will be followed up with a short telephone interview (up to 30 minutes), which you have booked.

If you have any questions or concerns, please email: Jessica.K.Baxendale@student.shu.ac.uk.

Best wishes,
Jess

Jessica Baxendale
PhD Student

Centre for Regional Economic and Social Research
Sheffield Hallam University
City Campus

Howard Street

Sheffield

S1 1WB

Phone: x

EMAIL TO BE SENT CONFIRMING PARTICIPATION

Dear [X]

Many thanks for agreeing to participate in the online workshop exploring evidence use and its role in decision making processes.

You will be sent an email inviting you to join the online workshop on Monday 25th June at 9am. You need to log in as soon as possible and register your details and view the other workshop participants. Then you will be able to log in and out whenever is convenient for you.

The workshop will be followed up with a short telephone interview (up to 30 minutes), which you have booked for [X].

If you have any questions or concerns, please email: Jessica.K.Baxendale@student.shu.ac.uk.

Best wishes,

Jess

Jessica Baxendale

PhD Student

Centre for Regional Economic and Social Research

Sheffield Hallam University

City Campus

Howard Street

Sheffield

S1 1WB

Phone: x

EMAIL TO BE SENT ON THE 25th June 9AM

Dear [X]

Many thanks for agreeing to participate in the online workshop exploring evidence use and its role in decision making processes.

You need to log in as soon as possible [HERE](#), register your details, upload a photo of yourself and view the other workshop participants. Then you will be able to log in and out whenever is convenient for you.

The workshop will be followed up with a short telephone interview (up to 30 minutes), which you have booked for [X].

If you have any questions or concerns, please email: Jessica.K.Baxendale@student.shu.ac.uk.

Best wishes,

Jess

Jessica Baxendale

PhD Student

Centre for Regional Economic and Social Research

Sheffield Hallam University

City Campus

Howard Street

Sheffield

S1 1WB

Phone: x

EMAIL TO BE SENT AFTER WORKSHOP

Dear [X]

Many thanks for participating in the online workshop exploring evidence use and its role in decision making processes.

I hope you found the experience interesting. If you have any further comments or feedback then please feel free to email me at Jessica.K.Baxendale@student.shu.ac.uk.

You can still log in and view/comment on the forum topics until Friday 12th July.

Best wishes,

Jess

Jessica Baxendale
PhD Student
Centre for Regional Economic and Social Research
Sheffield Hallam University
City Campus
Howard Street
Sheffield
S1 1WB
Phone: x

Appendix 13: Workshop participants

Table 1: Workshop participants

Name	Organisation	Role	Completed Survey?
Ulrike Hotopp	BIS	Chief Economist	
Christine John	BIS	Governance Advisor	
Cath Goodall	BIS	Enterprise and Economic Development Analysis	Y
Christopher Mee	BIS	Senior Policy Advisor	
Sue Bishop	BIS	Director, Services Industries, Strategic Trade Group at UK Trade & Investment	
Jonathan Turner	BIS	Economic Analysis	
Ian Macfarlane Mitchell	BIS	Operational Research Manager	
Phill Adams	CLG	Economic Policy Manager	Y
Paul Mooney	ONE	Chief Economist ONE	
Tim Pain	ONE	Head of Business Enterprise and Skills ONE	
Jo Povey	ONE	Head of Business Investment, Programming and Planning ONE	
Robin Beveridge	ONE	Head of Policy and Research ONE	Y
Jon Carling	ONE	Current Head of NERIP ONE	
Sarah Mcmillan	ONE	Strategic Policy Manager ONE	
Carol Candler	ONE	Research and Evaluation Manager ONE	Y
David Mell	ONE	Modelling Manager	
William Rossiter	emda	Head of Strategy and Regional Affairs EMDA	Y
Andrew Morgan	emda	Skills and Communities Director EMDA	
Miles Burger	emda	Evaluation Manager EMDA	

Appendix 14: Quantitative workshop results

Table 1: Ranking the interventions

Question 2: Rank the projects from 1 to 10 (1 = most favourable project to receive investment, 10 = least favourable project to receive investment).	1	2	3	4	5	6	7	8	9	10	Response Count
1. Digital House	7	1	3	1	1	2	1	0	0	1	17
2. Regional events	0	1	0	1	4	3	0	2	1	5	17
3. Atown market regeneration	0	3	2	2	3	1	2	3	1	0	17
4. Charity support	1	1	0	0	0	2	0	7	4	2	17
5. Atown railway station	0	1	3	1	3	3	0	1	2	3	17
6. Encouraging enterprise	2	1	2	2	3	2	4	1	0	0	17
7. Saving the planet	1	3	0	3	1	1	2	2	3	1	17
8. Internship for graduates	4	3	2	1	0	2	3	0	2	0	17
9. Innovating together	2	2	4	2	2	0	1	0	1	3	17
10. Opportunities for all	0	1	1	4	0	1	4	1	3	2	17

Table 2: Ranking the interventions (%)

Question 2: Rank the projects from 1 to 10 (1 = most favourable project to receive investment, 10 = least favourable project to receive investment).		1	2	3	4	5	6	7	8	9	10	Response Count
1. Digital House		6%	6%	0%	0%	0%	12%	0%	41%	24%	12%	17
2. Regional events		0%	6%	0%	6%	24%	18%	0%	12%	6%	29%	17
3. Atown market regeneration		0%	6%	6%	24%	0%	6%	24%	6%	18%	12%	17
4. Charity support		0%	6%	18%	6%	18%	18%	0%	6%	12%	18%	17
5. Atown railway station		6%	18%	0%	18%	6%	6%	12%	12%	18%	6%	17
6. Encouraging enterprise		0%	18%	12%	12%	18%	6%	12%	18%	6%	0%	17
7. Saving the planet		12%	12%	24%	12%	12%	0%	6%	0%	6%	18%	17
8. Internship for graduates		12%	6%	12%	12%	18%	12%	24%	6%	0%	0%	17
9. Innovating together		24%	18%	12%	6%	0%	12%	18%	0%	12%	0%	17
10. Opportunities for all		41%	6%	18%	6%	6%	12%	6%	0%	0%	6%	17

Table 3: Overall rank and score

Calculated rank	Weighted score
1. Digital House	13.1
8. Internship for graduates	11.6
6. Encouraging enterprise	10.6
9. Innovating together	10.5
3. Atown market regeneration	9.9
7. Saving the planet	9
5. Atown railway station	8.3
10. Opportunities for all	7.7
2. Regional events	6.8
4. Charity support	6

Table 4: Exploring the data synthesis process¹²⁰

Question 6(a): What level of importance did you give the different pieces of information?							
Answer Options	Most important	Very important	Quite important	Important	Not very important	Response Count	
Project name/description	2	2	2	3	8	17	
Local Authority	0	1	3	7	6	17	
Forecast investment 2012-2013	5	10	2	0	0	17	
Total GVA in £s (for investment 2012-13)	5	3	8	1	0	17	
Jobs created	4	10	3	0	0	17	
Businesses created	3	11	3	0	0	17	
Businesses supported	1	8	7	1	0	17	
People assisted to gain employment	2	6	7	2	0	17	
People assisted in skills	0	7	6	3	1	17	
CO2 reduced (tonnes)	0	4	5	5	3	17	
Political/pragmatic considerations	2	3	7	4	1	17	

¹²⁰ For reasons of clarity, in the narrative the results of ‘most important’ and ‘very important’ are combined to express agreement.

Table 5: Exploring the data synthesis process (%)

Question 6(a): What level of importance did you give the different pieces of information?						
Answer Options	Most important	Very important	Quite important	Important	Not very important	Response Count
Project name/description	12%	12%	12%	18%	47%	17
Local Authority	0%	6%	18%	41%	35%	17
Forecast investment 2012-2013	29%	59%	12%	0%	0%	17
Total GVA in £s (for investment 2012-13)	29%	18%	47%	6%	0%	17
Jobs created	24%	59%	18%	0%	0%	17
Businesses created	18%	65%	18%	0%	0%	17
Businesses supported	6%	47%	41%	6%	0%	17
People assisted to gain employment	12%	35%	41%	12%	0%	17
People assisted in skills	0%	41%	35%	18%	6%	17
CO2 reduced (tonnes)	0%	24%	29%	29%	18%	17
Political/pragmatic considerations	12%	18%	41%	24%	6%	17

Table 6: Context of the decision making process

Question 8(a): How easy/difficult do you think it would be to rank the projects if:				
Answer Options	Very easy/Easy	Neither easy or difficult	Very difficult/Difficult	Response Count
There was a clear strategy i.e. reduce unemployment	15	1	1	17
There was more project information given	10	4	3	17
A small team needed to come to a decision	6	7	4	17
The number of projects increased	1	5	11	17
A large team needed to come to a decision	0	2	15	17
There was no clear strategy	0	0	17	17

Table 7: Context of the decision making process (%)

Question 8(a): How easy/difficult do you think it would be to rank the projects if:				
Answer Options	Very easy/Easy	Neither easy or difficult	Very difficult/Difficult	Response Count
There was a clear strategy i.e. reduce unemployment	88%	6%	6%	17
There was more project information given	59%	24%	18%	17
A small team needed to come to a decision	35%	41%	24%	17
The number of projects increased	6%	29%	65%	17
A large team needed to come to a decision	0%	12%	88%	17
There was no clear strategy	0%	0%	100%	17

Table 8: Utilisation of decision support tools

Question 11(a): Would you use such a tool for decision making?			
Answer Options	Response Percent	Response Count	
Definitely would use	11%	2	
Likely to use	67%	12	
Not likely to use	22%	4	
Definitely wouldn't use	0%	0	

Table 9: Perceived benefits of decision support tools¹²¹

Question 12(a): What are the positives of using such decision tools?						
Answer Options	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree	Response Count
Assists collaborative decision making	3	10	4	1	0	18
Uses available data	2	16	0	0	0	18
Decision makers can work with analysts to ensure tool is appropriate	4	10	2	2	0	18
Enables decisions to be replicable	3	11	1	3	0	18
Enables decisions to be transparent	6	7	1	3	1	18
Enables total benefits and costs to be calculated	2	12	2	2	0	18
Flexibility if strategic/budget decisions change	0	14	1	1	1	17
Would enable better outcomes (GVA/outputs etc.)	0	6	9	3	0	18

¹²¹ For reasons of clarity, in the narrative the results for 'strongly agree' and 'agree' are combined to express agreement.

Table 10: Perceived benefits of decision support tools (%)

Question 12(a): What are the positives of using such decision tools?						
Answer Options	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree	Response Count
Assists collaborative decision making	17%	56%	22%	6%	0%	18
Uses available data	11%	89%	0%	0%	0%	18
Decision makers can work with analysts to ensure tool is appropriate	22%	56%	11%	11%	0%	18
Enables decisions to be replicable	17%	61%	6%	17%	0%	18
Enables decisions to be transparent	33%	39%	6%	17%	6%	18
Enables total benefits and costs to be calculated	11%	67%	11%	11%	0%	18
Flexibility if strategic/budget decisions change	0%	78%	6%	6%	6%	17
Would enable better outcomes (GVA/outputs etc.)	0%	33%	50%	17%	0%	18

Table 11: Perceived negatives of decision support tools¹²²

Question 13(a): What are the negatives of using such decision tools?						
Answer Options	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree	Response Count
Poor monitoring data	1	11	3	3	0	18
Poor evaluation data	2	12	3	1	0	18
Poor socio-economic data	1	7	6	3	0	17
Poor contracting cost data	1	7	7	3	0	18
Unable to capture all qualitative information	11	7	0	0	0	18
Hidden assumptions	4	10	1	0	1	16
Difficult to understand the tool	0	5	2	8	3	18
Potential mistakes in the programming of the tool	1	8	5	4	0	18
Would not enable better outcomes (GVA/outputs etc.)	0	0	14	4	0	18

¹²² For reasons of clarity, in the narrative the results for 'strongly agree' and 'agree' are combined to express agreement.

Table 12: Perceived negatives of decision support tools (%)

Question 13(a): What are the negatives of using such decision tools?						
Answer Options	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree	Response Count
Poor monitoring data	6%	61%	17%	17%	0%	18
Poor evaluation data	11%	67%	17%	6%	0%	18
Poor socio-economic data	6%	39%	33%	17%	0%	17
Poor contracting cost data	6%	39%	39%	17%	0%	18
Unable to capture all qualitative information	61%	39%	0%	0%	0%	18
Hidden assumptions	22%	56%	6%	0%	6%	16
Difficult to understand the tool	0%	28%	11%	44%	17%	18
Potential mistakes in the programming of the tool	6%	44%	28%	22%	0%	18
Would not enable better outcomes (GVA/outputs etc.)	0%	0%	78%	22%	0%	18

Appendix 15: Qualitative workshop data tables

Table 1: Workshop qualitative data tables

Theme	Sub-theme	Quotation	Survey question that generated the quotation ¹²³	Quotation used in thesis narrative?
Investment prioritisation without the use of a decision support tool (scenario 1): Strategic investment prioritisation in the RDAs	Strategic investment prioritisation in the RDAs	Ranking projects is always tricky, comparing apples and pears. Hence allocating budgets to programme areas, and then sub-programmes, and ranking within those sub programmes is easier - comparing apples with apples. However, the choice then has to be made in allocating levels of budget to different programmes and sub-programmes. (ONE officer [11])	8(b): How easy/difficult do you think it would be to rank the projects if: Please explain your decisions above	Y
		Individual projects cannot be judged against each other in my view unless they are similar in nature - this needs to be part of a strategic decision that says we need to spend XX on infrastructure, YY on skills and ZZ on business support to achieve a certain amount of GVA. (ONE officer [16])	11(b): Would you use such a tool for decision making? Please explain your decision above	Y
		Would look at projects within individual programme/sub themes, to compare similar projects to each other. (ONE officer [11])	11(b): Would you use such a tool for decision making? Please explain your decision above	Y
Investment	CBA approach	I prioritised jobs created and GVA and investment	6(b): What level of	Y

¹²³ See Appendix 7 for workshop items. See Appendix 2 for online survey items and respondents.

<p>prioritisation without the use of a decision support tool (scenario 1): A hypothetical investment decision making scenario</p>		costs to generate a benefit cost ratio which I feel is key to decision making. (ONE officer [7])	importance did you give the different pieces of information? Please explain your decisions above	
		Using a cost benefit analysis in which I'd optimise impact (jobs, GVA and benefit/cost ratio) within the budget constraint but [allow] some "wriggling" round the edges where... there was some other strategic factor to play in (e.g. political partnerships). (ONE officer [7])	5: In real life how would you have decided which projects should go forward?	N
		I used an approach that looked at the ratio of costs to deliverables. I took account of evaluation evidence for similar kinds of projects with which I am familiar. And in the context of the current economic context, prioritised job creation and business starts. (EMDA officer [1])	6(b): What level of importance did you give the different pieces of information? Please explain your decisions above	Y
		My reaction to Scenario 1 was to build into the spreadsheets a cost benefit model with weights for the benefits given. (Central Government officer [8])	10: How do you think this scenario compares to scenario 1?	Y
	Complex decision making behaviour	I gave each one a score for 'priority' based on my gut feel of the fit with my personal RDA priority list, another score for outputs, and a third for GVA: £1 ratio (noting...scepticism about the figures). I then ranked each of these scores and calculated an average weighting. Finally, I then fiddled the weightings until I got a result I was happy with!	Forum: How did you weigh up the different pieces of information to come to a decision?	Y

			(ONE officer [11]. Forum comment)	Forum: How did you weigh up the different pieces of information to come to a decision?	Y
			I ruled out some (ranked 5-10) due to lack of rationale... or expensive and no exit strategy... In real world projects would need a business case, with economists looking at economic case (I'm not an economist - as you can probably tell!). Subject to that I ranked the others on - policy imperative..., evidence of (likely) success...; vfm and in an area of known need for the economy (skills, enterprise). I looked at cost per job as a v. rough indicator of vfm. (Central Government officer [18]. Forum comment)		
			Balance between factors for decision making	16: From your experience, what do you think is the best way to package evidence to feed through into decision making processes for prioritising investment?	Y
			Need for a broad range of evidence	5: In real life how would you have decided which projects should go forward?	Y
Investment prioritisation without the use of a decision support tool (scenario 1): An identified role for knowledge translation tools		Complexity of the decision making environment	In my prioritisation I used an underlying clear strategy based on economics. Without this it would have been very difficult. With a large team it is more difficult to agree this underlying strategy and ensure everybody is pulling in the same direction. (Central Government officer [2])	8(b): How easy/difficult do you think it would be to rank the projects if: Please explain your decisions above	Y
			There was a tendency with RDA investment	16: From your experience,	Y

	decisions - that I saw - for the volume of evidence to stand in the way of clear decision making. (ONE officer [12])	what do you think is the best way to package evidence to feed through into decision making processes for prioritising investment?	
	Too much information can put decision makers off. Important to find ways of summarising lots of complex information clearly. (EMDA officer [9])	16: From your experience, what do you think is the best way to package evidence to feed through into decision making processes for prioritising investment?	Y
	The process must be sensitive to the limited capacity of decision makers to process information. (EMDA officer [1])	16: From your experience, what do you think is the best way to package evidence to feed through into decision making processes for prioritising investment?	Y
Strategic and cherry-picked evidence use	Personally I would have chosen those projects which deliver the highest forecast GVA and which deliver GVA over a long period of time (i.e. physical redevelopment). In reality, political considerations - both national and local - would inevitably mean that 'pet' projects would be championed and commissioned. (EMDA officer [5])	5: In real life how would you have decided which projects should go forward?	Y
	Political realities can mean that the most justifiable decision may be different from the best decision. For this reason it is often tricky to accurately record decision making processes. (ONE officer	9: How do you think the process of decision making can be made replicable and transparent?	Y

Using a decision tool to support an EBRPM approach (scenario 2): Receptivity to the decision support tool for investment prioritisation	Positive receptivity	[16]) Much better - more evidence-based, transparent and data-driven. (EMDA officer [5])	10: How do you think this scenario compares to scenario 1?	Y
		We should do much more of this in decision making. (EMDA officer [9])	14(b): Would such a decision tool have been useful for prioritising investment: Other & please explain your decisions above	Y
		I would do this - and did work up a similar model at ONE. It's an excellent way of actually using real evaluation data to help inform future investment decisions. (ONE officer [7])	11(b): Would you use such a tool for decision making? Please explain your decision above	Y
	Decision tool promoting the systematic utilisation of evidence	This sort of tool is a great way of using evaluation evidence intensively and effectively. (ONE officer [7])	12(b): What are the positives of using such decision tools? Other & please explain your decisions above	Y
		The tool provides much more information in a consistent format. The complexity and volume of information could itself become a problem for decision makers - M. Feldman 1989 and Kingdon 1995 have both written interestingly on this. I think the systematic use of evaluation evidence/benchmarks is a strength of the model. (EMDA officer [1])	10: How do you think this scenario compares to scenario 1?	Y
	Enabling data cleaning, synthesis and	It would ensure consistency in figurework and allow various financial options to be considered with relative ease, providing there is accuracy in	10: How do you think this scenario compares to scenario 1?	Y

		input data. (Central Government officer [3])		
		Much more nuanced analysis, using a complex and effective-seeming model. (ONE officer [12])	10: How do you think this scenario compares to scenario 1?	N
		Similar in many ways. But presentation of the data is much more transparent and easier to slice and dice. (ONE officer [7])	10: How do you think this scenario compares to scenario 1?	Y
		A very useful tool to aid decision making, by holding and ordering a lot of information in a single place as well as ensuring the process is transparent so all involved in the process are clear on the criteria. (ONE officer [6])	11(b): Would you use such a tool for decision making? Please explain your decision above	Y
	Address constraints to individuals' analytical processing of information Counter cherry picked and unsystematic evidence use	The tool helps inform the process, particularly with large numbers of projects. (ONE officer [10])	10: How do you think this scenario compares to scenario 1?	Y
		While no tool will capture all qualitative data, this nonetheless takes us forward. And it is an improvement on many of the finger in the air/ personality based investment decisions made in many organisations. (ONE officer [7])	13(b): What are the negatives of using such decision tools? Other & please explain your decisions above	Y
		Also very useful to be able to show a robust method for prioritising spend. (ONE officer [6])	11(b): Would you use such a tool for decision making? Please explain your decision above	Y
		This sort of tool is a great way of using evaluation evidence intensively and effectively. Main benefits	12(b): What are the positives of using such decision tools?	N

		are around this and the transparency of at least this part of decision making. (ONE officer [7])	Other & please explain your decisions above	
		Much more structured and transparent method of informing decision making. (ONE officer [6])	10: How do you think this scenario compares to scenario 1?	N
		[There needs to be] better recording and communications of decisions to enable a body of good practice to be built up. (Central Government officer [13])	15: From your experience, how do you think decision making processes for prioritising investment can be improved?	N
Using a decision tool to support an EBRPM approach (scenario 2): The role of a decision support tool to shape policy decisions	Complexity of decision making	There is a risk that this process can become over scientific. (EMDA officer [9])	10: How do you think this scenario compares to scenario 1?	Y
		Plugging figures into a sausage machine such as this discourages true, deeper analysis of the projects - e.g. use of BIS benchmarks of gross to net factors instead of examining the actual gross to net figures achieved by a particular project, which can vary very widely. (ONE officer [11])	12(b): What are the positives of using such decision tools? Other & please explain your decisions above	Y
		Danger of oversimplifying the process. (Central Government officer [4])	13(b): What are the negatives of using such decision tools? Other & please explain your decisions above	Y
	Political nature of evidence and decision making	You can't get away from the political angle, and maybe you shouldn't try to. Robust quantitative models are great, but you have to allow people to have their say otherwise the process looks remote and done by eggheads in ivory towers. (ONE officer	15: From your experience, how do you think decision making processes for prioritising investment can be improved?	Y

		[12])		
		There is a balance to be struck between what is pragmatic (i.e. public and political considerations) and the outcome of quantitative or other analysis. My top ten choices were made on the basis of the data provided, without spending as long as I should have on proper quantitative analysis. But in the real world, you have to be aware of the views of politicians and the electorate, and that could change the decisions made. (ONE officer [12])	6(b): What level of importance did you give the different pieces of information? Please explain your decision above	Y
	Decision tool to support dialogue within the wider policy process	It seems to me to form the basis of a dialogue - you can [show] stakeholders what the model says, and which projects it prioritises, and then have a discussion about the pros and cons. Much of this discussion will be of a political nature. (ONE officer [12])	11(b): Would you use such a tool for decision making? Please explain your decisions above	Y
		The data does need to be packaged through some kind of tool/ mechanism, but it needs to sit alongside discussion, negotiation and proper explanation. (ONE officer[16])	16: From your experience, what do you think is the best way to package evidence to feed through into decision making processes for prioritising investment?	N
		Would be an excellent tool to use in the initial scoping phase of a project, prior to it going to the Board for the political discussions to be had. (EMDA officer [5])	11(b): Would you use such a tool for decision making? Please explain your decisions above	N
		Use of such models to inform decisions made in	15: From your experience.	N

		light of wider considerations - but in knowledge of GVA, jobs, cost impacts etc. Face to face group discussions with various parties brought together helps surface and pursue issues, particularly with one or more parties offering challenge. (Central Government officer [18])	how do you think decision making processes for prioritising investment can be improved?	
		The outputs must match the decision makers' styles. Traffic light (Red Amber Greens) summarise views. (Central Government officer [8])	16: From your experience, what do you think is the best way to package evidence to feed through into decision making processes for prioritising investment?	N
	Data quality	Any tool which is data driven inevitably lives or dies by the data which drives it. (EMDA officer [5])	13(b): What are the negatives of using such decision tools? Other & please explain your decisions above	Y
		From experience, a lot of the core data is a bit iffy. So the tool is only as good as the data that's put in. Screening the poor data would need to be a key task upfront. (ONE officer [7])	13(b): What are the negatives of using such decision tools? Other & please explain your decisions above	
		The model is clearly a useful tool, but is clearly highly dependent on the quality of data involved. Having worked on the evaluation framework both in the RDA's and Government, that would be a key challenge. (Central Government officer [17])	10: How do you think this scenario compares to scenario 1?	Y
		If the quality and quantity of both evaluation AND monitoring data been improved, it might have been possible to set up systems to make	Survey question 21: How do you think evaluation evidence could/should have	Y

		performance (including return on investment and social and environmental performance) within and across projects more visible to the strategic decision makers. (RDA evaluation officer [59])	informed strategic decision making at the RDA?	
	The use of GVA	I tended not to believe the GVA figures. Outputs, evaluation findings and knowledge about the delivery partner are stronger factors, but harder to do objectively. (EMDA officer [9]. Forum comment)	Forum: How did you weigh up the different pieces of information to come to a decision?	Y
		Idiosyncratic activities may have little evaluation evidence on which to base any benchmarks. (EMDA officer [5])	13(b): What are the negatives of using such decision tools? Other & please explain your decisions above	Y
		Some of the assumptions are open to criticism. In particular, the multipliers from the PWC evaluation may be the best available, but they may vary from region to region or for other reasons. (ONE officer [12])	12(b): What are the positives of using such decision tools? Other & please explain your decisions above	Y
	Cost per output data	The flaw at [the RDA] was following through with hard metrics on cost per output. (RDA project delivery/performance officer [29]. Online survey)	Survey 52: In your opinion do you think conducting evaluations is good value for money? Thinking about your answer above, how do you think value for money can be improved?	Y
		It could give information in a much more tangible form of the various output/ outcome options	10: How do you think this scenario compares to	Y

		which can be of more importance than the actual financial reduction. (Central Government office [3])	scenario 1?	
	Evaluation data problems	Robustness of sample size was highly variable. Could only have been improved through longitudinal or a better approach to maintaining beneficiary data. (RDA evaluation officer [54]. Online survey)	Survey 59: How were robust sample sizes for primary data collection determined and quality checked at the RDA? How do you think this could have been improved?	N
		If the quality and quantity of both evaluation AND monitoring data been improved, it might have been possible to set up systems to make performance (including return on investment and social and environmental performance) within and across projects more visible to the strategic decision makers. They would also need to be very aware of the limitations of the data used including the levels of confidence they could be expected to have in them. Then they might have felt more comfortable using evaluation to its potential in strategic decision making. (RDA evaluation officer [58]. Online survey)	Survey 21: How do you think evaluation evidence could/should have informed strategic decision making at the RDA?	Y
		Need to take into account risk. Sensitivity analysis should be more common. (Central Government officer [4])	15: From your experience, how do you think decision making processes for prioritising investment can be improved?	Y
		Know from experience that evaluation data from different sources may be 'calibrated' differently.	13(b): What are the negatives of using such decision tools?	Y

	(ONE officer [10])	Other & please explain your decisions above	
Monitoring data problems	To clarify - re lack of monitoring data and its robustness. This was due to staff not inputting outputs onto the computer system in a timely fashion and by staff not being professional when they inputted it and essentially making it up at times. (RDA strategy officer [63]. Online survey)	Survey 27(b): How important were each of the following challenges for the RDA when implementing the budget cuts in 2010? Other (please specify) / further comments	Y
	Simply having these systems in place was not enough as the approach to monitoring was not consistent across all teams. (RDA delivery/performance officer [34]. Online survey)	Survey 43(b): How would you rate the robustness of the RDA's processes for monitoring? Other (please specify) / further comments	Y
	The whole system of monitoring outputs was very complicated though, and people's interpretation of the guidance on what constituted a verifiable output made monitoring really difficult. (RDA delivery/performance officer [12]. Online survey)	Survey 41(b): Was monitoring data on spend, outputs and beneficiary data used to inform project delivery at the RDA? If yes, how?	Y
	Using cleansed data will provide a useful tool and could avoid costly mistakes going forward. (Central Government officer [3])	13(b): What are the negatives of using such decision tools? Other & please explain your decisions above	Y
Underlying logic model	As for monitoring and evaluation data, the trick is to establish what you need to collect BEFORE you start commissioning projects. (EMDA officer [5].	Forum: How do you think the accuracy of monitoring data, socio-economic data, and	Y

		Forum comment)	evaluation data can be improved?	
The construction of a decision support tool: Employing a decision logic	Key issue was the valuation and direct comparison of outputs	Context is the big issue - context, mechanism and outcomes are rarely differentiated. (RDA evaluation officer [55]. Online survey)	Survey 15: What do you think are the main challenges for synthesising the current evaluation evidence base to inform decision making for the LEPs and the Regional Growth Fund? Other (please specify) / further comments	Y
		More use of recognised logic chains. (Central Government officer [3])	9: How do you think the process of decision making can be made replicable and transparent?	Y
		My main concern was to focus on the core business of an RDA, and away from areas that are other organisations' responsibilities, such as CO2 reduction or skills. (ONE officer [11])	6(b): What level of importance did you give the different pieces of information? Please explain your decisions above	Y
		I would prioritise the regeneration and business creation schemes, and leave the less tangible programmes. Experience and hindsight suggests they are the scheme that got the greatest bang for your money. (Central Government officer [17]) Co2 measures are often underrated but shouldn't be. (ONE officer [7])	5: In real life how would you have decided which projects should go forward?	Y
			6(b): What level of importance did you give the	Y

Developing the decision support tool	Nuanced approach needed to understand decision-maker's preferences			different pieces of information? Please explain your decisions above	
		Confidence in the data. Stakeholder agreed criteria, scorings and weightings. (Central Government officer [8])	7(b): What factors do you think are important when coming to a decision? Please list any other factors you think are important	Y	
		I think such a model is very useful but it would need modification and the decision maker should be able to add some additional knowledge into the decision making process. (Central Government officer)	11(b): Would you use such a tool for decision making? Please explain your decision above	Y	
	Developing the decision knowledge base: tacit knowledge	Getting the decision makers to articulate their views would be the key to this exercise. Without this process it will be difficult to get a basis for decision that will bear scrutiny. (Central Government officer [8])	8(b): How easy/difficult do you think it would be to rank the projects if: Please explain your decisions above	Y	
		I don't think that the model provides adequate coverage of more qualitative sources of information that are often very important in the context of investment decisions. (EMDA officer [1])	12(b): What are the positives of using such decision tools? Other & please explain your decisions above	Y	
		It is important to provide space in the appraisal process for more qualitative inputs - particularly practitioner's experience of delivery. (EMDA officer [1])	10: How do you think this scenario compares to scenario 1?	N	

		This model would be a massive improvement. Also need to include realistic assessment of likely outputs and more evaluation and lessons learned information. (EMDA officer [9])	15: From your experience, how do you think decision making processes for prioritising investment can be improved?	N
	Robust quantitative data and context of data	Numbers are not the full story, and neither should they be. But as a first step they provide an excellent means of sorting the wheat from the chaff. (EMDA officer [5]. Forum comment)	Forum: How do you think the accuracy of monitoring data, and socio-economic data, and evaluation data can be improved?	Y
		The reality is usually that you need to delve beneath the figures to understand the true impact of the project and the need for the funding. (ONE officer [16])	6(b): What level of importance did you give the different pieces of information? Please explain your decisions above	Y
	Additional factors to consider in the tool	By the end of [the RDA], we would have used an appraisal process that included use of evaluation evidence to inform decisions. However, more qualitative inputs such as views on deliverability, risks associated with the project and 'political' sensitivities were also taken into account. (EMDA officer [1])	5: In real life how would you have decided which projects should go forward?	Y
		It takes into account the same kind of factors that I was considering when making judgements in scenario 1. It is a more thorough and rigorous approach than the one I took. It doesn't, as far as I can tell, allow for flexibility to deal with political or	10: How do you think this scenario compares to scenario 1?	Y

		pragmatic issues. It also doesn't seem to factor in the power of negotiation with project partners - a project might not be worth going ahead with at the current cost, but at a reduced cost may well be worth it. (Central Government officer [15])		
		Considered deliverability, outputs and impact. Would want to consider long term legacy of investment, while also tackling short term problems, like youth unemployment and low skill levels. Existing commitments cost of existing planned schemes and political fallout are also very important. Evaluation impact also should be a factor to enable the wider impact (outcomes) to be a factor in decisions. (EMDA officer [9])	5: In real life how would you have decided which projects should go forward?	N
		I would use it as part of the process, understanding there are quite often other factors which also need to be taken into account in decision making, i.e. practicalities, politics, partners, short term priorities, etc. (ONE officer [6])	11(b): Would you use such a tool for decision making? Please explain your decision above	Y
		Also need to include realistic assessment of likely outputs and more evaluation and lessons learned information. (EMDA officer [9])	15: From your experience, how do you think decision making processes for prioritising investment can be improved?	N
	Experiential knowledge	As ever the decisions are only as good as the underlying information available. The danger is that decision making is over-dependent on just cost-benefit analysis and does not provide enough	12(b): What are the positives of using such decision tools? Other & please explain your decisions above	

		possibilities for well-informed experts with experience of appraising and making investment decisions.(Central Government officer [4])		
		There is a danger that in reality a system like this would add to an already bureaucratic process of decision making. It would need to replace current systems, but also allow some flexibility to consider other factors. (ONE officer [16])	11(b): Would you use such a tool for decision making? Please explain your decision above	N
		Considered deliverability, outputs and impact. Would want to consider long term legacy of investment, while also tackling short term problems, like youth unemployment and low skill levels. Existing commitments cost of existing planned schemes and political fallout are also very important. Evaluation impact also should be a factor to enable the wider impact (outcomes) to be a factor in decisions. (EMDA officer [9])	5: In real life how would you have decided which projects should go forward?	N
		There are other factors not considered, such as timing, genuine need for the grant, and more detail behind SAV... I'm not sure how the systems consider issues such as local impact? (ONE officer [16])	11(b): Would you use such a tool for decision making? Please explain your decision above	N
		The quality of outcomes or quantity of outputs are usually down to the people delivering and managing the projects. (Central Government officer [3])	12(b): What are the positives of using such decision tools? Other & please explain your decisions above	N
		Level of investment and political/pragmatic considerations reflect the reality of the situation.	6(b): What level of importance did you give the	N

		(EMDA officer [9])	different pieces of information? Please explain your decisions above	
		I think it is important to provide space in the appraisal process for more qualitative inputs - particularly practitioners' experience of delivery. (EMDA officer [1])	10: How do you think this scenario compares to scenario 1?	N
	Value judgements	Outputs important, but gut instinct about the types of projects that are most successful is important but hard to measure. (EMDA officer [9])	6(b): What level of importance did you give the different pieces of information? Please explain your decisions above	Y
		The tool would have been helpful to inform the decision making process, but would have presumably only been part of the process. I don't think it can replicate the personal judgment of the project officer e.g. whether a project could be scaled back and still achieve sufficient outcomes. (Central Government officer [15])	14(b): Would such a decision tool have been useful for prioritising investment: Other & please explain your decisions above	Y
		While it is important to capture the information that scenario 2 provides, and important to use this sort of model to defend difficult decisions, the risk is that it produces outcomes that "feel" wrong. (EMDA officer [9])	10: How do you think this scenario compares to scenario 1?	Y
	Perception that qualitative evidence	Model is only as good as the data that goes into it. Impossible to create a model that incorporates	13(b): What are the negatives of using such decision tools?	Y

	could not be incorporated appropriately into decision tools	qualitative data. (ONE officer [16])	Other & please explain your decisions above	
Developing the decision support tool: Multi Criteria Analysis	Multi Criteria Analysis	<p>Decide on the criteria for assessing bids, such as: Rationale for govt intervention, economic impact, jobs created (should be part of economic impact - the above GVA data didn't look like it was consistent), political priorities. Use simple multi-criteria analysis to rank bids and adjust using experience in assessing bids to come up with a list of suggested funding options. Give senior decision makers (e.g. ministers) decision on final projects.</p> <p>In real life and, if asked for advice, I would have recommended some form of MCD (multi-criteria decision making) technique. Such techniques, which are usually applied in a workshop setting and require good quality facilitation, aim to (i) explicitly tease out the criteria against which decisions are being made (ii) attempt to build some consensus and understanding about the trade-offs between different decision options in terms of these criteria (iii) place the type of information that former colleagues have identified (the details that might emerge in negotiation and project shaping) in a coherent framework. These techniques are ideal for choosing between a range of options particularly where resource constraints</p>	5: In real life how would you have decided which projects should go forward?	Y
			Forum: In real life how would you have decided which projects should go forward?	Y

	exist. They help identify the strengths and weaknesses of particular options and, through this, can lead to option improvement as a side-product. (ONE officer [19]. Forum quotation)		
	My preference would be to see a matrix system for quantitative assessment alongside a qualitative matrix based on evaluation material, experience of the subject and knowledge of the sector/ delivery agent. This would feed into a priority matrix (against government policy) and used to determine what appears to be best value for money. A group of expert advisors would then give their opinion on the proposal to limit prejudice and to achieve a consensus. The differing viewpoints add value to the likely success of the project. (Central Government officer [3]. Forum quotation)	Forum: In real life how would you have decided which projects should go forward?	N
	A combination of methods has worked effectively from 2001 to 2006 to address this question for defence equipment. This combined systems thinking with simulation [to] provide descriptive and predictive models of a process [and] to capture expert knowledge. The predictive runs of a simulation [model] allow what-if and combinations of policy interventions to be understood. The utility of these outputs will vary according to the decision makers' and stakeholders' perspectives. Multi Criteria Decision analysis used their arguments to provide a prescriptive analysis to	15: From your experience, how do you think decision making processes for prioritising investment can be improved?	N

Developing the decision support tool: Updating data over time	Updating data over time	decide what options should be pursued. (Central Government officer [8])	13(b): What are the negatives of using such decision tools? Other & please explain your decisions above	Y
		The underlying assumptions would need to be regularly updated or re-considered. (EMDA officer [5])	Forum: How do you think the accuracy of monitoring data, socio-economic data, and evaluation data can be improved?	N
		The trick with data which is used for decision making is to ensure that it is relevant, up-to-date and transparent (in its assumptions and how it is collected). (EMDA officer [5]. Forum comment)	13(b): What are the negatives of using such decision tools? Other & please explain your decisions above	Y
	Updating data over time in decision tool	To be reasonably accurate the data needs to be captured early in the projects life cycle and continually updated as retrospection can have its own flaws. (Central Government officer [3]) This requirement for evaluation should be across all projects, regardless of size and complexity, and should be based on a singular set of guidelines (so the same data is being collected for all projects). This means that at the end of a project it is possible to say what the actual benefits of the work have been (and using the concept of your model, these findings would then be incorporated into the benchmarks figures used to assess future projects). (EMDA officer [5]) There needs to be a well-defined feedback mechanism whereby the results of the evaluation feed into future similar work (in the case of	16: From your experience, what do you think is the best way to package evidence to feed through into decision making processes for prioritising investment?	N

Developing the decision support tool: Programmer skill and the need for capacity building		scenario 2, this would mean that the various benchmarks would need to be updated regularly). (EMDA officer [5]. Forum comment)	evaluation data can be improved?	
	Assumptions	Decision making is easier [with the tool] but may be over-reliant on the underlying data and assumptions. (Central Government officer [4])	10: How do you think this scenario compares to scenario 1?	Y
		Too many assumptions made, giving an illusion of objectivity whereas really the subjective decisions have been hidden in the assumptions built into the model. (ONE officer [11])	11(b): Would you use such a tool for decision making? Please explain your decision above	Y
	Need to pilot test	I would want to see it applied to some real life examples to judge its effectiveness. (Central Government officer [15])	11(b): Would you use such a tool for decision making? Please explain your decision above	Y
		Piloting some of the ideas to establish if the options will work as expected seems critical to their utility. (Central Government officer [18])	14(b): Would such a decision tool have been useful for prioritising investment: Other & please explain your decisions above	N
		There would need to be a control test against historic projects (or similar) to validate the model and the data. (Central Government officer [15])	13(b): What are the negatives of using such decision tools? Other & please explain your decisions above	N
Need to understand the limitations of data and analysis		While it gives an apparent objective comparison between projects, it is hard for non-experts to see how the numbers have been arrived at. (ONE officer [11])	12(b): What are the positives of using such decision tools? Other & please explain your decisions above	Y
		Clear indication of the assumptions and evidence	15: From your experience,	Y

		used to calculate the cost and benefits. (Central Government officer [4])	how do you think decision making processes for prioritising investment can be improved?	
	Capacity building	<p>The biggest problem which such a tool is that it would need to be both endorsed and ratified at the highest political level (whether it be at an internal, local, regional or national level). (EMDA officer [5]. Forum comment)</p> <p>Tool isn't too difficult to understand but there would have to be a very good understanding, and buy-in, across the organisation so it had credibility. (ONE officer [6])</p> <p>The public sector, alas, has many barriers which inhibit [decision support tools] use including significant cultural factors and a lack of analytical understanding. (ONE officer [19]. Forum quotation)</p> <p>Decision makers, who include non-economists - would need to have confidence in the tool and the data. (Central Government officer [18])</p> <p>A model that has the trust and confidence of the decision makers could be a powerful tool. (Central Government officer [15])</p>	<p>Forum: Could such models be applied to decision making in the future?</p> <p>13(b): What are the negatives of using such decision tools? Other & please explain your decisions above</p> <p>Forum: In real life how would you have decided which projects should go forward?</p> <p>13(b): What are the negatives of using such decision tools? Other & please explain your decisions above</p> <p>16: From your experience, what do you think is the best way to package evidence to feed through into decision making processes for prioritising investment?</p>	<p>Y</p> <p>Y</p> <p>Y</p> <p>N</p> <p>N</p>

		<p>It needs to be a careful balance between evidence and common sense. Both the RDAs and Government have a tendency to swing between evidence based decision making and gut feeling too often (see the Regional Growth Fund at the moment). A steady middle ground would improve replicability and consistency. (Central Government officer [17])</p>	<p>16: From your experience, what do you think is the best way to package evidence to feed through into decision making processes for prioritising investment?</p>	N
		<p>[There needs to be] learning and development about investment planning and knowledge of some basic concepts around net impact. That includes local politicians. Too many are blinded by the gross numbers on how many are helped (voters?) rather than whether that public spending actually made a difference in net terms - transparency of decision making is really important. (ONE officer [7])</p>	<p>15: From your experience, how do you think decision making processes for prioritising investment can be improved?</p>	N
		<p>There's a process of education about the benefits of using evidence effectively. Making sure people understand its purpose, pitfalls and potential is the first step. Thinking about creating a market for evidence is crucial and then working through issues of accountability. RDAs for example were forced - some reluctantly - to invest in evidence and evaluation. This was a really good thing. The fact that we have so much evidence to feed into the sort of model you've showcased is testimony to the power of coercion! Perhaps that degree of</p>	<p>16: From your experience, what do you think is the best way to package evidence to feed through into decision making processes for prioritising investment?</p>	N

		<p>accountability for public spending could be applied across the public sector more consistently and vigorously. Some central government departments do loads. Others don't. And most local authorities don't even know what evaluation is and don't invest in anything other than headcounts. (ONE officer [7])</p>		
--	--	---	--	--